OPTIMAL DESIGN AND OPERATION OF LIVESTOCK BREEDING PROGRAMMES WITH RESTRICTIONS IN INBREEDING

Santiago Avendaño

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

University of Edinburgh

2003

DEDICATION

I dedicate this thesis to three persons who unfortunately did no live to see me completing this challenge.

My grandfather Raúl (*Tata*), who seeded my primary interest in agriculture from early years in my life.

Mr. Carlos Luce, who was key in stimulating me to take this way.

My beloved grandmother Berta (*Nana*) who knew from the beginning that at some point I was going to fly abroad and from whom I have learnt how to observe life from different perspectives. Although we couldn't manage to get together once more, I know we are both sharing this. I also know that you have been next to me throughout this challenge.

DECLARATION

I declare that this thesis is my own composition and that the research described in it is my own work, except where otherwise stated.

Santiago Avendaño, 2003

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I believe that we are the only responsible and architects of our own lives and that every achievement rely on our own commitment and conviction, and in our ability to take opportunities when they appear. I also believe that for meeting our targets support in all aspects of life is essential. I have indeed received support from many people and institutions throughout these years. Now is time to express my gratitude.

I am deeply indebted to Beatriz Villanueva and John Woolliams for guiding me with endless patience and providing me with all kinds of support at all times during the course of my PhD. I have been very lucky in having such a good team of supervisors from whom I have learnt not only a few things on a particular subject, but also, and more importantly, the meaning of scientific excellence. At all times, I felt that I was part of a team rather than being merely supervised, and I am very grateful for their confidence. Working with both of them has been a great honour and I have enjoyed this time enormously. I also thank my 'silent' supervisor Peter Visscher for his sharp and fresh view on the research in this thesis. Ricardo Pong-Wong is acknowledged for patiently helping me with programming tasks in Fortran 90, something I enjoy, but am not naturally born for.

This work would not have been possible without the funding of the Meat and Livestock Commission through the LINK-SLP Programme (Grant LK0640). I also thank the National Institute of Agricultural Research (INIA) and the Agronomy Faculty of the National University in Uruguay for providing essential support to undertake the MSc/Diploma in Quantitative Genetics and Genome Analysis at Edinburgh University during 1999, an stepping point for this PhD project. I also thank the Scottish Agricultural College (SAC), in particular the Animal Breeding and Development team at Bush Estate for providing an excellent working environment.

I am afraid I have not enough words of gratitude for Geoff Simm. Without his constant support and encouragement since I arrived in Edinburgh in 1999 (and before!), this somewhat utopia of studying abroad would have never had happened. I

am also indebted to Marie Haskell, along with Geoff, for their warm friendship and hospitality during my first year in Edinburgh. They successfully managed to survive to several months of *uruguayan* idiosyncrasy and *latin* challenges that Carla and I put on their way. I am infinitely indebted to both of you.

I have met lots of very interesting people during these years who have contributed to many aspects of my everyday life. I wish to thank the people at SAC in the Stephen Watson Building, in particular Gillian, Jo, Spiridoula and Jos, Jim, Tim, Eileen, Elly, Ilias, Susan and Rick, Lutz, Anne and the group of sweet ladies that clean the building in the evening.

I have harvested an excellent crop of friends during these years, with whom I have shared all sort of good moments and adventures. Beatriz and Mike have been always there and I thank their warm friendship and essential support, particularly while trying to settle down here in Edinburgh during the Master's time. Mike has been a '*partner for everything*', we shared our passionate devotion to all sorts of meats, wine and football. We managed not only to have our unique all year round indoor barbecue, but also slaughtered and eaten our own heifer, '*orphan Annie*'. The *latin* contingent has been also key for keeping my spirit in good shape: Xulio and Carolina, Juan and Verónica, Manuel, Demian, Andrea, Rafael and Claudia, Emilio and Eduardo.

I must profoundly thank people in Uruguay who has backed me at all times before and while I was trying to achieve this. Ruben Severino has been my mentor in every sense, professionally and most importantly on a personal level, constantly stimulating me on every challenge. I will be eternally indebted to you. Thanks to Gabriel Rovere for being such a good friend and for advising me well. Diego Gimeno, my first academic mentor and friend is acknowledged for laying down good principles in animal breeding. Gonzalo Gonzalez and Jorge Urioste are greatly acknowledged for sorting out vital funding that allowed me to start this journey in 1999.

I must dedicate a special line to Alejandro Lombardi (*Pájaro*) and Cecilia Fernández. *Pájaro* is my brother from life, partner and permanent accomplice. We shared all sort of dreams and utopias, most of them we now see becoming true, while for the rest is just a matter of time. Cecilia was a sort of guardian angel that suddenly appeared, thanks for so much love and for being always there.

My sincerest and deepest thanks to my family for their love and support. I thank my mother in law Lyrian for helping me with my fatherhood duties during key times while trying to put together this thesis. My mother Beatriz has been battling alongside me during the last stretch of the writing up, witnessing my struggle and always supporting me with her endless energy and determination. Mum has been constantly stimulating and, indeed sponsoring, me since my University times, I know how important this is for you.

My final words are for Carla my wife who has been my walking stick during all these years, listening and backing me during struggling times and sharing my successes. You have made my life infinitely rich with your love, freshness and commitment. The completion of this PhD thesis was our shared objective and is also your achievement. Pablo our baby son has also done his bit, being patient and keen on sleeping when I cuddled him while typing at the same time. Carlita, we are now ready to start the next stage of our enjoyable life together.

Yes I know, a little bit long list of acknowledgements, but as I said at the beginning, this was the time to be grateful.

PUBLICATIONS

Avendaño, S., Villanueva, B., and Woolliams, J.A., 2002. Expected increases in genetic merit in the UK Aberdeen Angus beef cattle breed from applying optimised selection. In: *Proceedings of the British Society of Animal Science*, 8-10 April 2002, University of York, York, p.54.

Avendaño, S., Villanueva, B. and Woolliams, J.A., 2002. Optimisation of selection decisions in the UK Meatlinc breed of sheep. *Proceedings of the British Society of Animal Science,* 8-10 April 2002, University of York, York, p.194.

Avendaño, S., Visscher, P.M. and Villanueva, B., 2002. Potential benefits of using identified major genes in two trait Breeding goals under truncation and optimal selection. In: *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production,* Montpellier, France. CD-ROM communication no. 23-03.

Avendaño, S., Villanueva, B. and Woolliams, J.A., 2003. Expected increases in genetic merit from using optimised selection in the Aberdeen Angus pedigree breed in the UK. In: *British Cattle Breeders Club, Preparing for the future, Annual Conference Papers*, Digest 58: 18-21.

Avendaño, S., Woolliams, J.A. and Villanueva, B., 2003. Predicting genetic gain when rates of inbreeding are constrained to pre-defined values. *Proceedings of the British Society of Animal Science,* 24-26 March, 2003, University of York, York, p.46.

Avendaño S., Villanueva, B., and Woolliams, J.A., 2003. Expected increases in genetic merit from using optimised contributions in two livestock populations of beef cattle and sheep. *Journal of Animal Science* (Accepted).

Avendaño S., Woolliams, J.A., and Villanueva, B., 2003. Mendelian sampling term as a selective advantage in optimum breeding schemes with restrictions on the rate of inbreeding. *Genetical Research, Cambridge* (Submitted).

ABSTRACT

Modern breeding programmes of livestock species have successfully led to increased genetic merit in traits of economic relevance through accurate and intense selection. However, concomitant increased levels of inbreeding have been also observed. Quadratic optimisation constitutes a general approach to the joint management of the rates of genetic gain (ΔG) and inbreeding (ΔF) in selected populations. The rate of inbreeding can be used as a measure of risk in the breeding programme. The method optimises the genetic contributions of selection candidates for maximising ΔG while restricting ΔF to a pre-defined value. The ΔF restriction is achieved by applying a quadratic constraint on the average co-ancestry of selection candidates weighted by their projected use. The general objectives of this thesis were: i) to implement and evaluate the potential benefits of quadratic optimisation in real livestock populations; ii) to develop a deterministic framework for predicting ΔG under constrained ΔF and iii) to evaluate the benefits of quadratic optimisation in multiple trait scenarios under mixed inheritance models.

The application of quadratic optimisation in two populations of beef cattle (Aberdeen Angus) and sheep (Meatlinc) led to important increases in the expected ΔG. At the observed ΔF in each population, increments per year in ΔG of 17% for Meatlinc and 30% for Aberdeen Angus were found in comparison to the ΔG expected from conventional truncation BLUP selection. More relaxed constraints on ΔF allowed even higher increases in expected ΔG in both populations.

Stochastic simulations have revealed that under quadratic optimisation the selective advantage of the candidates for selection is primarily their Mendelian sampling terms rather than their breeding values as under truncation selection. Thus, under quadratic optimisation, the contribution of candidates to the future genetic pool is decided upon the best information on their unique superiority or inferiority with respect to the parental mean.

A self-contained and accurate deterministic approach for predicting ΔG for pre-defined ΔF has been developed. It requires only specification of the trait heritability, the number of selection candidates and the target ΔF.

Benefits from quadratic optimisation were also evaluated in a two-trait scenario where the trait with lower heritability was affected by an identified quantitative trait loci (QTL). Extra gains in the breeding goal were observed throughout the whole selection process from the combined use of both optimised contributions and QTL information. In contrast, this scheme was not the most effective for improving each of the traits in the breeding objective.

The design and operational tools developed in this thesis constitute a general framework for the evaluation and realisation of the benefits from quadratic optimisation tools in practical livestock breeding programmes.

1. CHAPTER ONE

General Introduction

1.1. Introduction

In this General Introduction current approaches for the management of inbreeding in livestock breeding populations are outlined. The aim is to place the research chapters of this thesis in the appropriate methodological context and to highlight the areas that required further development before this thesis was started. Also, the specific research objectives and deliverables of this research are specified.

1.2. Inbreeding as a component of risk in livestock breeding programmes

Whilst inbreeding is unavoidable in closed selection programmes, the increases in inbreeding need to be restricted to alleviate its negative effects. Inbreeding promotes the reduction of the genetic variance in the selected trait, hence imposing theoretical limits to selection, and induces random changes in gene frequency of both neutral and selected loci (Falconer and Mackay, 1996; Caballero and Santiago, 1998). Increased rates of inbreeding decrease the probability of fixation of rare alleles with selective advantage in the selection scheme, and can lead to more dramatic changes in the frequency of undesirable alleles (Caballero and Santiago, 1998). Also, when gene effects are dominant inbreeding is expected to lead to phenotypic depression in selected production traits and in other traits not included in the breeding objective, particularly in fitness related traits (Falconer and Mackay, 1996; Lynch and Walsh, 1998). Keller *et al.* (1990) using a deterministic model for closed beef nucleus breeding schemes predicted a much greater impact of inbreeding depression on reducing genetic response than the effects of inbreeding on genetic variance.

There is documented evidence of increases in inbreeding coefficient (F) and inbreeding depression in production traits in most livestock species. Burrow (1993) reviewed the effects of inbreeding in beef cattle, finding negative effects in a broad number of economically important traits. Consistently, growth traits from birth to maturity were affected by inbreeding and showed reductions of 0.06, 0.44, 0.69 and

1.30 kg in live weight at birth, weaning and maturity, respectively for each percent increase in F in the individual. Importantly, inbreeding impacted all measures of heifer fertility. Overall, the effects of inbreeding on performance were more pronounced in young animals, when animals were inbred at higher rates, and when F was higher than 20%.

Lamberson and Thomas (1984) summarised the effects of inbreeding in sheep populations and also found detrimental effects in performance for most economically important traits. Fitness related traits such as ewe fertility (i.e., ewes lambing per ewe joined) and lamb survival (i.e., lambs weaned per lambs born) were particularly affected and the estimated reduction was 0.014 and 0.028, respectively for each percent of increase in individual F. These values represented 1.6% and 3.3% reductions from the corresponding mean performance in a reference flock for each trait, respectively. In an experimental flock, Wiener *et al.* (1994) summarised the results of their previous studies on the effect of inbreeding depression on different aspects of sheep production and found that rapid increases in inbreeding had a high impact on the component traits of overall productivity. The flock gross income per kilogram of metabolic live weight of ewe was reduced by about 1.2% per each percent increase in F. These authors pointed out that deleterious effects of this magnitude are also expected in commercial flocks where lower inbreeding levels are typically observed. This was because of the largely linear observed effects of inbreeding on financial returns for a range of F up to 50% and because for the component traits, the observed inbreeding depression was consistent with the range of that reviewed by Lamberson and Thomas (1984), which involved studies at lower inbreeding levels.

In dairy cattle, recent estimates in the United States (US) Holstein by Thompson *et al.* (2000) indicate significant reductions in milk production (from 35 up to 708 kg per each percentage increase in F for the range of individual F from 2% to 6% and over 20%, respectively). These reductions represent around 0.3% and 7.0% per percent F (for a reference 305-day lactation of 10,000 kg). The effects of F were greatest early in the lactation and in young animals and for F values greater than 10%

impacted age at calving (increased in up to 26 days) and in lactation length (decreased up to 8 days). Inbreeding had a detrimental effect on survival for all levels of F and for all lactations and was identified as the potential greatest negative effect on economic loss. Smith *et al.* (1998) estimated inbreeding depression on total lifetime performance (measured as relative net income) of 6% of the additive standard deviation for milk production and days of productive life. Inbreeding levels are currently increasing in dairy cattle populations. Weigel (2001) estimated the current rate of inbreeding of the US populations of Ayrshire, Brown Swiss, Guernsey, Holstein and Jersey to be 0.3%, 0.8%, 0.8%, 1.3% and 1.6% per generation, respectively. Weigel (2001) quoted that some individuals have sired up to 250,000 milking daughters and 3,000 progeny tested sons, and that about 50% of young Holstein bulls tested globally are offspring of only 10 sires.

In pigs, Rodrigañez *et al.* (1998) reported mean reductions of 0.27 piglets born and 0.39 born alive for each 10% of litter F in Large White. These would correspond to an inbreeding depression of about 0.3% and 0.5% per percent increase in F, respectively assuming an average litter size of eight piglets. Fernández *et al.* (2002) found performance reductions from the mean ranging from 0.3% to 0.5% for 120 days weight, and from 0.3% to 0.6% for daily gain for each percent increase in F in three strains of Iberian pigs.

Woolliams *et al.* (2002) have proposed the rate of inbreeding (ΔF) as a design parameter and as a measure of risk from the perspective of the breeding programme. This is justified by the three following reasons. First, ΔF or equivalently, the effective population size (i.e., $N_e=1/[2\Delta F]$) is the relevant population parameter that determines the level of standing genetic neutral variability and the effects of genetic drift on neutral loci (Falconer and Mackay, 1996; Hill, 2000; Hill, 2003). Secondly, and more relevant for the time span involved in livestock breeding programmes, population processes like the rate of fixation or loss of favourable mutations and selection against deleterious mutations (e.g., Caballero and Santiago, 1998; Hill, 2000) are expressed in terms of ΔF. Thirdly, ΔF is linked to the effects of drift on the variance of response (Falconer, 1996; Nicholas, 1987). For instance, desirable levels

of ΔF can lead to acceptable levels of variation in response, particularly when generation intervals are short and/or selection is intense and accurate (Meuwissen and Woolliams, 1994b).

Modern breeding programmes of livestock species have been successful in improving the genetic merit of those traits of economic relevance. The development of selection criteria with increasing amounts of family information like family indices and Best Linear Unbiased Prediction (BLUP) of breeding values (EBV) allows higher selection accuracy thus enhancing the possibility of increased genetic gains in economically important traits. On the other hand, higher concomitant levels of inbreeding are expected when using these selection criteria due to the increased probability of co-selection of relatives, as the weight given to the family information is higher. This could be particularly important in reproductive and fitness related traits since index theory indicates that more emphasis is placed on family information when the heritability is low.

The studies of Quinton *et al.* (1992) and Verrier *et al.* (1993) have been illustrative on the consequences in the long-term of the exclusive use of BLUP-EBVs. These authors showed that when compared at the same level of inbreeding, less accurate methods like mass selection can give higher gains than BLUP, and that at least in small populations, the short-term benefits in response from increased accuracy from using BLUP selection may be offset in the medium or long term. BLUP is nowadays the standard method for genetic evaluation in livestock breeding programmes. For instance, in the United Kingdom (UK), BLUP-EBV from multi-trait genetic evaluations have been available since the early nineteen nineties for beef cattle and sheep (Wray and Simm, 1991).

Modern breeding programmes have also been ready to develop and utilise reproductive technologies that allow additional genetic gains through higher selection intensities. However, the use of these technologies also lead to higher inbreeding rates. On the male side, AI (artificial insemination) has became a standard reproductive practice in livestock breeding programmes (e.g., beef and dairy cattle,

pigs, poultry and sheep). For instance, it is estimated that a fifth of the female world cattle population is currently bred by artificial insemination (Thibier and Wagner, 2002). On the female side MOET (multiple ovulation and embryo transfer) has had some recent uptake in dairy and beef cattle and in sheep. Villanueva *et al.* (1995) found cumulative gains from MOET to be about 50% higher than those from more traditional schemes, but ΔF was increased up to threefold when methods of controlling inbreeding were not used. Recently, van Arendonk and Bijma (2002) predicted similar benefits in rates of gain and concomitant increases in ΔF from MOET.

1.3. The joint management of inbreeding and gain in livestock breeding programmes

In the past, the development of strategies for controlling inbreeding levels in closed animal populations was mainly focused on small breeding populations in which conservation of genetic variation was the main objective (e.g., Caballero and Toro, 2000). In contrast, selection in livestock breeding programmes of domestic species has been traditionally concentrated on breeding objectives mainly focused on maximising the improvement of production traits. As mentioned before, accurate selection from BLUP-EBV and intense selection accompanied by increasingly efficient reproductive techniques has enabled important rates of genetic gain in livestock species. At the same time, concerns regarding increasing levels of inbreeding in commercial selected populations and its negative consequences on genetic variability and inbreeding depression in selected and non-selected traits are increasing (e.g., Notter, 1999; Weigel, 2001).

Hence, the development of methods for controlling inbreeding is relevant not only for conservation of populations of small size, but also for designing breeding programmes of livestock species with more commercial objectives towards the improvement of traits related to commercial products (e.g., meat, milk, eggs, fibre, etc.). It should also be emphasised that the prime objective pursued by commercial

breeding programmes is the achievement of the highest genetic responses at all times. Therefore, the trade off between seeking more accurate selection in the anticipation of additional genetic gain and the potential consequences of more rapid inbreeding becomes a relevant issue.

The joint management of genetic gain and inbreeding in livestock breeding programmes undergoing selection has been the subject of much recent research, and the main strategies developed in recent years will be outlined in this General Introduction. Methods for dealing with gain and inbreeding simultaneously can be grouped as in a decision making step process. Villanueva et al (1996) suggested a convenient break-up of tools with respect to the two decision stages arising when dealing with breeding schemes with restrictions on inbreeding. The first problem is the *a priori* design of the breeding scheme for a given set of design variables: resources (e.g., number of candidates per generation), trait heritability (h^2), and time scale. Tools for solving this decision stage will be referred to as 'Design tools'. The second problem is *a posteriori*, and refers to selection methods to be applied during the course of operation of the breeding scheme, and are consequently referred to as 'Operational tools'. Furthermore, two main decisions for the genetic management of the breeding scheme are required (e.g., Caballero and Toro, 2000): i) which animals should be selected and how much they should contribute to the future genetic pool; and ii) the mating policy to be applied among selected candidates.

Following the sequence of research chapters in this thesis, Operational tools will be presented first, followed by Design tools.

1.3.1. Operational Tools

1.3.1.1. The separate management of inbreeding and genetic response

Methods for controlling inbreeding in selected populations have often been developed with the objective of restricting or minimising the amount of inbreeding while the effects on genetic gain were considered separately.

At the selection stage, strategies range from simply increasing the proportion of selected candidates to limiting the emphasis given to the family information included in the selection criteria, or to limiting the number of selected individuals from a particular family. Toro and Perez-Enciso (1990) applied constraints on the number of selected individuals contributed by different families, thus exploring the range from within-family selection (Demple, 1975) through to family selection. Limiting the relative weight given to the family mean in the selection criteria was managed by Toro and Perez-Enciso (1990) when selection is based on a family index (i.e., including individual and family mean deviation terms) by setting a range of lower index weights than that which maximised selection accuracy. With BLUP selection, the relative weight given to the family information has been managed by the use of a biased (upward) heritability in the genetic evaluation (Grundy *et al.*, 1994), by selecting on modified indices that modify the relative weighing to the sire and dam component of the candidate EBV (Verrier *et al.*, 1993; Villanueva *et al.*, 1994) or by putting a cost factor on the individual inbreeding (Villanueva *et al.*, 1994).

Specific mating systems for reducing inbreeding such as minimum coancestry matings (Toro *et al.*, 1988), compensatory matings (Santiago and Caballero, 1995) or factorial matings (Woolliams, 1989) have also been implemented after the selection stage (Toro and Perez-Enciso 1990; Grundy *et al.*, 1994; Villanueva *et al.*, 1994). Selection and mating strategies have also been simultaneously considered following a mate selection strategy (Toro and Perez-Enciso, 1990). Even without being attached to a specific selection policy with regard to inbreeding (i.e., applying no restrictions in the optimal use of available information for the genetic evaluation), mating strategies on their own constitute a valuable tool for controlling rates of inbreeding (see Caballero *et al.*, 1996, for a review). These authors found up to a 25% reduction in ΔF when combining minimum coancestry and compensatory matings compared to random mating for both phenotypic and BLUP selection.

A general conclusion from these approaches is that although they can be successful in reducing the rate of inbreeding, they also generally lead to lower rate of gain (ΔG) than that achievable when inbreeding is not controlled. Although in some cases losses in genetic gain might be small or negligible (e.g., Villanueva *et al.*, 1994), methods that consider gain and inbreeding separately are expected to be sub-optimal with respect to gain.

All the above methods for controlling inbreeding were based on truncation selection, that is, only individuals above a certain EBV threshold were selected and equal usage was allocated to them, thus family sizes were fixed across generations. Toro and Nieto (1984) proposed an innovative approach for maximising the effective population size while restricting the selection differential (and therefore selection response) to the corresponding value in truncation selection. This was achieved by allowing differential usage of selected candidates. Using a quadratic programming approach, they minimised the sum of squares of contributions ($\sum c_i^2$) of selected candidates to the following generation, (hence, maximising [1] $N_e^* = 1/\sum c_i^2$) subject to the restriction $\sum c_i x_i = s$, where x_i is the estimated breeding value of the candidate i, and s is the selection differential under truncation selection. Although, more animals were selected, the desired selection intensity was achieved while more contributions were allocated to higher ranked individuals. This key feature of the so called weighted selection method constitutes one of the core concepts underlying the development of quadratic indices which will be the focus in the next section and throughout the thesis.

1.3.1.2. The simultaneous management of inbreeding and genetic response

---

[1] Although Toro and Nieto (1984) referred to $1/\sum c_i^2$ for the maximisation of N_e, this is related to the long-term genetic contributions (Wray and Thompson, 1990), rather than to initial contributions. Hence, N_e^* has been used instead.

Selection strategies for managing simultaneously genetic gain and inbreeding have been developed in recent years. The core objective was the achievement of an optimal solution for selection decisions that maximises genetic response with a restriction on the increase of inbreeding. The analytical way of making this objective explicit was by defining target objective functions in which both expressions for gain and inbreeding are included.

In broader terms the optimisation of breeding programmes can be seen as the maximisation of genetic gains while managing risk factors. The risk of the scheme can be summarised in different ways, for instance, by the variance of response ($V(\Delta G)$) or by ΔF, such that schemes with high rates of gain and low risks may be preferred over schemes with very high gains and high risks (Woolliams and Meuwissen, 1993; Meuwissen and Woolliams, 1994b). Meuwissen and Woolliams (1994b) emphasised the strong link between both risk components, since $V(\Delta G) \approx 2\Delta F \sigma^2_{A,\infty}(1-r^2_\infty k)$, where $\sigma^2_{A,\infty}$ is the equilibrium additive genetic variance at the Bulmer equilibrium, r^2_∞ is the equilibrium squared selection accuracy and k is the variance reduction coefficient (Bulmer, 1971).

Quadratic optimisation constitutes a general framework for the development of operational selection tools that jointly manage gain and risk. The idea of quadratic optimisation for the management of risk in breeding schemes was laid down by Meuwissen and Woolliams (1993) when developing a selection rule that maximised a function aimed to restrict the variance of selection response. The basic form of the function to be maximised was $\mathbf{c}^T\mathbf{g} - \lambda\mathbf{c}^T\mathbf{E}\mathbf{c}$, where $\mathbf{c}$ is the vector of contributions of candidates to the next generation, $\mathbf{g}$ is the vector of EBV, $\mathbf{E}$ is the matrix of prediction error (co) variances of EBV, and λ is a cost factor. Meuwissen and Woolliams (1994b) related the management of the variance of response and the management of inbreeding by noting that $\mathbf{E} = (\mathbf{Z}^T\mathbf{R}^{-1}\mathbf{Z} + \mathbf{A}^{-1}\sigma_A^{-2})^{-1}$, where $\mathbf{Z}$ is the design matrix relating phenotypic records to animals, $\mathbf{R}$ is the (co) variance matrix of environmental effects and $\mathbf{A}$ is the numerator relationship matrix. In the long-term, $\mathbf{E}$ tends to $\mathbf{A}\sigma^2_A$ (Mewissen and Woolliams, 1994b; Woolliams *et al.*, 2002). Since with

fully random union of gametes the term $\mathbf{c}^{\mathrm{T}}\mathbf{Ac}$ represents twice the inbreeding in the next generation, when the objective is to maximise genetic gain while constraining inbreeding, the function to be maximised can be expressed as $\mathbf{c}^{\mathrm{T}}\mathbf{g} - \lambda\mathbf{c}^{\mathrm{T}}\mathbf{Ac}$ (Woolliams *et al.*, 2002).

Therefore, as pointed out by Woolliams *et al.* (2002) the problem of short and long-term optimisation are intrinsically linked. In the short-term, a constraint on the quadratic expression $\mathbf{c}^{\mathrm{T}}\mathbf{Ec}$ (i.e., conditional on both records and pedigree) allows the management of the variance of response, while in the long-term a constraint on the quadratic term $\mathbf{c}^{\mathrm{T}}\mathbf{Ac}$ (i.e.,conditional to pedigree only) for **E** allows the management of inbreeding. The cost factor λ can be set arbitrarily or implemented as a Lagrange multiplier chosen to constrain the variance of response or the increase in inbreeding to pre-defined values. Both target functions using either **E** or **A** are subject to the constraints $0 \leq c_{i_x} \leq 0.5$, where c_{i_x} is the proportion of matings allocated to a selected candidate *i* of sex *x* such as $\sum c_{i_{males}} + \sum c_{i_{females}} = 1.0$.

A key property of quadratic optimisation functions is that they keep the method used for genetic evaluation and the restriction on risk (e.g., ΔF) separated. Therefore, EBVs can be estimated with the best available technique (i.e., BLUP) and the constraints independently chosen according to a preferred risk policy. This was not possible in previous strategies for controlling inbreeding that either restricted family sizes, limited the amounts of records included in the breeding value estimation or used biased genetic parameters for reducing the relative weight given to the family information.

Wray and Goddard (1994), Brisbane and Gibson (1995), Meuwissen (1997) and Grundy *et al.* (1998a) implemented dynamic selection algorithms in the form of quadratic indices of the form $\mathbf{c}^{\mathrm{T}}\mathbf{g} - \lambda\mathbf{c}^{\mathrm{T}}\mathbf{Ac}$ that constrained the weighed average relationship among selection candidates, hence the future inbreeding. The definition of the constraint has been a matter of attention as it affects whether the cumulative inbreeding or the rate of inbreeding are constrained.

Wray and Goddard (1994) and Brisbane and Gibson (1995) implemented constraints on the cumulative inbreeding coefficient rather than on ΔF. Wray and Goddard (1994) linked the restriction (Q) to the effects of inbreeding depression (D) on genetic gain such as $Q = \Delta G_L (t-1)/2 + D$, where ΔG_L is a prediction of the asymptotic rate of gain per generation in an infinite population including the effects of selection of genetic variance (Bulmer, 1971) and t is the time horizon. This definition of Q assumes a selection decision as, for predicting ΔG_L, predictions of the selection accuracy, the asymptotic genetic variance and the selection intensity are required. Thus, as pointed out by the authors, Q can only be approximated as the breeding goal and the cost imposed by inbreeding are not independent. The value of Q affects the selection decisions and the selection decision change the optimum Q value. In contrast, Brisbane and Gibson (1995) set a range of arbitrary constraint (k) values (e.g., from 0.3 to 10) to obtain different levels of genetic gain and cumulative inbreeding. Both approaches based the finding of the global optimum by iterative evaluations of the changes in the objective function after replacing individuals from the selected and unselected groups. Also, they assumed equal contributions by each selected candidate. Only in one of the strategies evaluated by Wray and Goddard (1994) not only the number of sires, but also the number of matings per sire was allowed to vary, providing extra benefits in gain. These authors envisaged the finding of an explicit maximum by using Lagrangian multipliers and differentiating the target function with respect to the usage of selected candidates.

The above restrictions on the absolute inbreeding coefficient are linked to the definition of a base population, which is hypothetical and often artificially defined. By restricting the absolute inbreeding coefficient, selection decisions will depend on the definition of the base population which should be not relevant to the current selection and breeding opportunities. Also, by defining the constraint in terms of inbreeding depression (e.g., Wray and Goddard, 1994) only one of the detrimental effects of inbreeding is accounted for. This not only could lead to constraints less stringent than those needed for instance to avoid the increase in gene frequency of deleterious genes but also its use in traits with complete additive effects is unclear. In

contrast, as it was mentioned previously, ΔF is the relevant parameter for summarising the long-term risk attached to the breeding programme.

The dynamic selection algorithms of Meuwissen (1997) and Grundy *et al.* (1998a) provide a solution for the problem of obtaining the optimum (i.e., maximum) rate of genetic gain while constraining the rate of inbreeding to pre-defined levels. The algorithm maximised the following objective function (Meuwissen, 1997):

$$H_t = \mathbf{c}_t^T \mathbf{g}_t - \lambda_0(\mathbf{c}_t^T \mathbf{A}_t \mathbf{c}_t - C_t) - [\mathbf{c}_t^T \mathbf{Q} - (1/2)\mathbf{1}^T]\boldsymbol{\lambda}$$

where $\mathbf{c}_t$ is the solution vector of mating proportions (c) of candidates at generation t, $\mathbf{g}_t$ is the vector of EBV of selection candidates, $\mathbf{A}_t$ is the numerator relationship matrix for selection candidates, $\mathbf{Q}$ is a known incidence matrix for the sex of the candidates, $\mathbf{1}^T$ equals [1 1], and λ_0 and $\boldsymbol{\lambda}$ are Lagrangian multipliers. The desired value for the inbreeding rate in the long-term is achieved by setting the constraint of Grundy et al (1998a); i.e., $C_t = 2[1-(1-\Delta F)^t]$. For a single generation the constraint is achieved by setting C_t to $2[\Delta F + (1 - \Delta F)F_t]$, where F_t is the average inbreeding coefficient of selection candidates. The third term in the objective function ensures that male and female parents contribute with a half of the gene pool each.

With this algorithm, an explicit solution is found at the selection stage of the breeding scheme in terms of both the optimal number of selected candidates and their optimal contributions to the next generation. Selected candidates are those with $c > 0$ and will contribute accordingly to their c value. A key feature of this dynamic selection algorithm is that for a given constraint, greater contributions will be allocated to individuals with greater EBV. Also, as the constraint imposed is less severe, fewer individuals will be selected, the usage of the individuals will became more unequal, and the intercept and slope of the regression of EBV on usage will be greater (Grundy *et al.*, 1998a). This was, the underlying idea of the weighted selection method of Toro and Nieto (1984).

Importantly, the selection decisions are conditional on the genetic relationships between the candidates and on their EBV. In contrast, under standard truncation selection, candidates are selected based on their EBV and independently from their relationships and equal usage is allocated to each selected individual.

Meuwissen (1997, Appendix) also provide extensions to the algorithm to accommodate extra constraints that may come from reproductive limits (e.g., a fixed female contribution or a maximum contribution per male) and extensions have been developed for overlapping generations (Sonesson and Meuwissen, 1998; Grundy *et al.*, 2000).

The benefits of this dynamic selection algorithm over truncation selection have only been evaluated under simulation. At the same ΔF, the extra ΔG from quadratic over truncation selection ranged from 21 to 60% under discrete generations (Meuwissen, 1997), and were up to 44% (Meuwissen and Sonesson, 1998) and 35% under overlapping generations (Grundy *et al.*, 2000). Mating systems can also be incorporated after the selection step adding further benefits. When compared to random mating, minimum coancestry matings with a maximum of one offspring per mating pair allowed extra genetic response from 5 to 23% under discrete generations and from 11 to 18% under overlapping generations (Sonesson and Meuwissen, 2000 and 2002, respectively). Consistently, the extra benefits from mating systems tend to be greater for stringent ΔF constraints and small scheme sizes. These extra benefits arise from homogenising relationships across families reducing the relationship among individuals with high EBV, thus giving more scope to quadratic optimisation for selecting animals with the highest EBV. Also, avoiding extreme relationships in the offspring and parents of the next generation, leads to a larger Mendelian sampling variance promoting increased genetic variance and genetic gain (Sonesson and Meuwissen, 2002).

The framework of quadratic indices is not restricted to polygenic inheritance models only and the dynamic selection tools of Meuwissen (1997) and Grundy *et al.* (1998a) have also been used in mixed inheritance models where the total breeding value has a

polygenic and quantitative trait loci (QTL) component. Villanueva *et al.* (1999) showed under BLUP selection the potential benefits from combining the selection on an identified QTL (GAS) and optimal contributions in reducing and even avoiding the long-term losses in genetic gain often reported when selecting on a QTL (e.g., Gibson, 1994; Pong-Wong and Woolliams, 1998). Furthermore, when optimal weights for the polygenic and QTL components in the total breeding value are combined with quadratic optimisation, benefits are consistently expected both in the short- and long-term (Villanueva *et al.*, 2002a). Relevantly, the use of the quadratic index (i.e., optimised contributions) only, without considering the QTL information achieved greater response than standard truncation selection (i.e., fixed contributions) using the genotype information across all generations. Thus, the greater impact on gains arose from the optimisation of contributions while the optimisation of weights to the polygenic and QTL components had a higher impact in allowing sustained gains in the long-term (Villanueva *et al.*, 2002a). Benefits under mixed inheritance models from quadratic optimisation are also expected when selection is on linked markers (MAS) rather than on the QTL itself (Villanueva *et al.*, 2002b). On the other hand, the marker distance to the QTL is not the relevant parameter that determines the degree of realisation of the expected upper bound benefits from quadratic optimisation under GAS. Even for a very tight linked marker (e.g., 0.05 cM) MAS achieved half of the gains obtained with GAS and the maximum gain was delayed one generation. However, when extra independent prior information about the QTL effects was supplied to the genetic evaluation, MAS was able to capitalise on the expected benefits from GAS (Villanueva *et al.*, 2002b).

Quadratic indices have potential not only in situations where the main objective is to achieve the maximum genetic gain while constraining ΔF to a pre-defined value according to a certain risk policy, like in commercial breeding programmes. The benefits of quadratic indices in the context of conservation schemes have also been evaluated. Eding *et al.* (2002) implemented a quadratic index to decide the optimal contributions of candidate populations to a gene bank that maximised the core set kinship. Sonesson *et al.* (2003) showed the potential role of quadratic indices for restricting or minimising ΔF in conservation schemes when a selection policy against

a disease gene is implemented. Also, Villanueva *et al.* (2003) reformulated the problem for the minimisation of ΔF (i.e., minimisation of $\mathbf{c}^{\mathrm{T}}\mathbf{A}\mathbf{c}$) for a pre-defined level of genetic gain (i.e., $\mathbf{c}^{\mathrm{T}}\mathbf{g} \geq K$) and derived a selection algorithm similar to that of Meuwissen (1997). The key concept is that with this approach both selection and conservation objectives can be seen as extremes of a broad optimisation problem in which different relative emphasis is given to gain and inbreeding.

Although simulation studies are consistent about the potential extra benefits from quadratic optimisation over traditional truncation selection, published information on the practical implementation in real livestock population is scarce. To date, the study of Weigel and Lin (2002) has been the only one that applied the dynamic selection tool of Meuwissen (1997) in livestock populations (i.e., dairy cattle), but concluding that genetic gain may be compromised by applying constraints in inbreeding. Although, these authors evaluated the gain obtainable for a broad range of inbreeding constraints, they did not compare the gain expected from the optimisation to that obtained with the current selection methods at the same rate of inbreeding, thus failing to establish a fair comparison. Moreover, they applied a direct constraint on the absolute level of inbreeding in the next generation, which did not guarantee a constant rate on a long-term basis.

Operational tools in the form of dynamic selection algorithms for the implementation of quadratic optimisation with restrictions on ΔF have been improved, become increasingly sophisticated and have the potential for being implemented in a wide range of scenarios. On the other hand, the benefits over standard truncation selection have been evaluated only in simulation studies that have the intrinsic limitation of being able to explore only a restricted number of scenarios in terms of parameter values. In addition, to date, it is not clear to what extent the anticipated benefits from optimising contributions using quadratic indices are realisable when compared to current selection strategies (e.g., truncation) at the same ΔF.

1.3.2. Design tools

Design tools are developed to provide *a priori* answers on the optimal use of available resources in terms of the size and structure of the breeding programme (e.g., mating ratios and number of offspring per parent). As explained before with operational tools, the objective is to maximise a function to achieve optimum genetic response while constraining the risk according to a particular policy. A distinctive feature of design tools is that deterministic predictions for the rate of gain and the components of risk (i.e., increase in inbreeding, or the variance of response) are required.

Goddard and Smith (1990) optimised the effective number of bulls for maximising net response in economic merit, accounting for the effects of inbreeding. The objective function maximised each generation t was $\Delta G_t - D\Delta F_t$ where D is the depression in economic merit per unit of inbreeding such as $D = d\sigma / CV$ where d is the percentage depression per percent of inbreeding, and σ and CV are the standard deviation and coefficient of variation of economic merit, respectively. No effects of selection on genetic variance were included as it was assumed that selection was practised in previous generations and an equilibrium was reached. They evaluated a range of number of bulls tested per generation (T) from 50 to 10,000 and three values of D (0.25, 0.5 and 1.0% per percent of inbreeding). It is worth noting that D = 1.0% is similar to the value found experimentally by Weiner *et al.* (1994) of 1.2% decrease in overall gross income per increase in F in sheep previously presented. Goddard and Smith (1990) found that the optimum number of bulls increased with the scheme size and with the D level. For instance, for D = 1.0% the optimum increased from 9 to 20 bulls when T increased from 50 to 10,000, respectively. For T = 10,000 the optimum number of bulls increased from 7 to 20 when D was increased from 0.25% to 1.0%. Interestingly, they found flat optima for the number of bull sires that gave the maximum net response at the three values of inbreeding depression in economic merit. In most cases between 8 to 12 bull sires per generation give more than 95% of the maximum net response. They also suggested a minimum of 10 bull sires per

generation (i.e., $N_e = 40$) for the whole Holstein breed that will give rise to ΔF of 1.25% per generation, a figure that they judged acceptable.

Goddard (1992) argued that the net gain approach ($\Delta G_t - D\Delta F_t$) represents a short-term criterion. For this reason he modified the function and used a net present value (NPV) approach for evaluating the short- and long- term gains and losses from selection and inbreeding depression, respectively. Further, he added also the effects of inbreeding on reducing genetic variance and the long-term contribution of mutational variance. The optimum number of bulls to maximise NPV was substantially greater than those to maximise net gain, hence leading to lower levels of inbreeding. The author concluded that from a global perspective the maximisation of NPV constitutes a safest approach than maximising net gain each generation. As it was argued in the case of the operational tool of Wray and Goddard (1994) the cost of inbreeding is mainly represented by the effects of inbreeding depression, thus restricting the approach to traits in which directional dominance is important. For instance, when Goddard (1992) set the value of D to 0, the optimum number of bulls was 2, thus implying no consideration of risk components in the design of the breeding scheme.

Meuwissen and Woolliams (1994c) developed a design tool for maximising genetic gain while constraining the variance of response in dairy breeding schemes. The optimum number of selected candidates per breeding path and generation intervals were found by maximising the function $\Omega = -E(\Delta G) + k(V(\Delta G) - CV^2E^2(\Delta G))$ where $E(\Delta G)$ and $V(\Delta G)$ are the expectation and variance of genetic gain, respectively and k is a large positive constant for obtaining coefficients of variation of response lower than a critical value (CV_c). Deterministic procedures were developed for calculating $E(\Delta G)$ and $V(\Delta G)$ and simulated annealing algorithms were used for solving the combinatorial problem of allocating animals to breeding paths and age classes. Maximised genetic gains were not sensitive to changes in the constraint on the coefficient of variation of annual response. Genetic gains were only slightly reduced from 0.30 to 0.29 genetic standard deviations when CV_c was restricted to 0.32 or 0.16, respectively. Breeding schemes changed considerably as

the constraint was more stringent, leading to optimal breeding schemes with formal progeny testing (rather than to schemes relying on young bulls from a nucleus), increased number of selected animals and openness of the scheme. In addition, coefficients of variations were more sensitive to deviations from optimal schemes than the rates of gain.

Deterministic optimisation tools for maximising the rate of gain while explicitly constraining the rate of inbreeding have been also developed. Under truncation selection, breeding programmes with constraints on ΔF can be optimally designed for mass selection (Villanueva *et al.*, 1996 and 2000 for discrete and overlapping generations, respectively) and index selection (Villanueva and Woolliams, 1997 for discrete generations). These authors derived asymptotic expressions for ΔG accounting for the effects of linkage disequilibrium (i.e., Bulmer effect) and inbreeding on genetic variance, and based their predictions of ΔF on the approaches of Wray *et al.* (1990) and Woolliams and Thompson (1994). A single linear objective function was maximised $\Phi = \Delta G - \lambda \Delta F$, where λ is a weighting factor taking positive values and increased at appropriate steps until the constraint in ΔF was achieved. Villanueva *et al.* (1996) also implemented this approach to restrict the variance of response by substituting ΔF for CV in the target function. The maximisation of this target function provided the optimum number of sires and dams to be selected given a fixed number of candidates and trait h^2 and allowed optimisations at specific points in time. In the case of index selection not only the number selected, but also the weights given to family information were optimised. Villanueva and Woolliams (1997) found that the superiority of index selection over mass selection depended upon the time horizon, population size, h^2, and the ΔF constraint. With restricted ΔF and optimised index weights the advantage of index over mass selection was greater as the time horizon was expanded. Optimised index selection had sizeable advantages (e.g., over 5%) over mass selection only for low h^2 and large scheme sizes (more index weights placed on family information) and for high h^2 and small scheme sizes (less index weights placed on family information). Also, the optimum selection proportions for maximising gain at the selection plateau

were around 2/3 for males and females, hence, greater than the figure of 1/2 proposed by Robertson (1960).

In the case of overlapping generations, both the number of males and females to be selected and their distribution across age classes are to be optimised. Villanueva *et al.* (2000) found that as the restriction on ΔF is more severe, the optimal strategy for maximising gain is different from truncation selection across age classes. Also, the restriction on ΔF was mainly achieved by increasing the number of candidates selected and to a lesser extent by increasing the generation interval. Only in small schemes severe restrictions led to increased generation intervals.

These deterministic studies provided practical insight on the optimum design of breeding schemes with constrained ΔF. A general finding was that the maximum ΔG was achieved at optimum mating ratios of one for small schemes, low h^2 and severe restrictions, whereas the optimal mating ratio increased up to two as the restriction was relaxed and the scale of the scheme increased.

The novelty of the approach of Villanueva *et al.* (2000) was that although they used the same target function $\Phi = \Delta G - \lambda \Delta F$, the predictions of ΔG and ΔF were based on expressions in term of the 'unified theory of long-term genetic contributions' put forward by Woolliams and Thompson (1994b).

The long-term contribution r of an ancestor i born at time t_1 is defined as the proportion of genes from i that are present in individuals born in generation t_2 (descendants) which derive by descent from i, with the condition that enough generations separate t_1 and t_2 (Wray and Thompson, 1990; Woolliams *et al.*, 1999). Thus, long-term contributions represent a measure of the proportional contribution of an individual to the current genetic make-up of the population. Over many generations, in a population thoroughly mixed, the r of an ancestor will converge to the same value for all its descendants but will differ among ancestors (Woolliams *et al.*, 1999).

Two basic expressions relating r to the rates of gain and inbreeding have provided a relevant framework for the development of deterministic optimisation of breeding schemes. First, Wray and Thompson (1990) demonstrated that the rate of inbreeding is proportional to the sum of squares of r, $E[\Delta F] = \frac{1}{4}\sum r^2$ and secondly, Woolliams and Thompson (1994) defined the rate of gain in terms of the covariance between r and the Mendelian sampling terms (a) as $E[\Delta G] = \sum r_i a_i$. Thus, both the genetic size of the population and the process of genetic improvement can be described in terms of long-term contributions (i.e., the individual gene flow) from ancestors to descendants. In particular, Woolliams and Thompson (1994) and Woolliams *et al.* (1999) emphasised that this definition makes explicit that sustained genetic gain relies on the creation of a covariance between r and a, thus better ancestors contribute more genes to future generations, and on the use of the newly created genetic variation each generation. Relevantly, Woolliams *et al.* (1999) emphasised that $E[\Delta G] = \sum r_i a_i$ also makes explicit the links between genetic response and the pedigree, not as evident in $\Delta G = i\rho\sigma_A$ (i.e., the 'breeder's equation').

Under truncation selection, Woolliams *et al.* (1999) and Woolliams and Bijma (2000) showed that the expected values for the asymptotic rate of gain and inbreeding can be written in terms of expected long-term contributions as $E[\Delta G] = \sum E[u_i a_i]$ and $E[\Delta F] = \frac{1}{2}\sum E[u_i^2]$, respectively, were u_i is the expected lifetime contribution of the individual i. The basic framework for predicting u_i was laid down by Woolliams *et al.* (1999) and is based on $u_i = E(r_i \mid s_i)$ where s_i is the selective advantage of the individual i. The expected long-term contribution of an ancestor i in category q (male or female) was given by $u_i = \alpha_q + \beta_j(s_i - \bar{s}_q)$, where α_q is the expected contribution of an average parent in category q (i.e., sex or age classes) and β_j is the regression of the contribution of the individual i on its selective advantage (expressed as a deviation from selected contemporaries in the same category $\bar{s}_q$).

The prediction framework for u_i has been the key development for the prediction of ΔF in selected populations. The key component of this derivation is that under truncation selection and for a given selective advantage (s), the error variance of r_i ($\sigma^2_{r_i}$) is proportional to the square of its mean ($\mu^2_{r_i}$), having the relevant consequence that for predicting ΔF, modelling u_i was only required (Woolliams and Bijma, 2000). Predictions of u_i have been developed for truncation schemes under mass or sib indices (Bijma and Woolliams, 1999) and BLUP (Bijma and Woolliams, 2000) selection. In particular, the derivation of explicit predictions of ΔF has constituted the key feature for the deterministic design of livestock breeding schemes under truncation selection and BLUP evaluation (Bijma *et al.*, 2001).

This deterministic prediction framework based on long-term contributions is now available for its use in real livestock breeding programmes under truncation selection and where candidates are selected based on mass, index or BLUP selection and for discrete or overlapping generations. The software Sel*A*ction (Bijma and Rutten, 2002; Rutten *et al.*, 2002) provides deterministic predictions of ΔG based on the pseudo-BLUP method of Wray and Hill (1989) and ΔF which is predicted upon the theory of long-term contributions. Predictions are performed after specification of sets of design variables, including genetic and phenotypic parameters and the population structure (e.g., number of selected candidates per sex, number of offspring per dam, etc).

In contrast to the advances in predictive approaches under truncation selection mentioned above, a deterministic framework under quadratic indices, that is, for predicting the maximisation of ΔG for a given ΔF is constrained to pre-defined levels is not available. This clearly requires the prediction of ΔG in schemes using dynamic selection algorithms with constrained inbreeding (Meuwissen, 1997; Grundy *et al.*, 1998a). Woolliams and Thompson (1994) suggested that the problem of maximising genetic gains for constrained ΔF could be solved by linearly allocating long-term contributions and Mendelian sampling term of selected candidates, hence relating strategic optimisation to their definition of ΔG. Grundy *et al.* (1998a) provided a

theoretical proof for this, and developed a deterministic prediction for the theoretical optimum ΔG for a given ΔF after specification of three design variables: the number of candidates (T), the trait h^2, and the desired ΔF. Their simulation study was conclusive in that the observed rates of gain under quadratic optimisation were always lower than the theoretical expectation under an exact linear allocation of long-term contributions and Mendelian sampling terms. The two reasons for this inability of attaining the optimum are: first, at the time of selection long-term contributions of candidates cannot be set to their expected values independently from previous generations; and secondly, Mendelian sampling terms are not known, thus the selection and usage of individuals can only be decided upon initial estimates of their unique superiority (or inferiority).

Therefore, there is an existing gap in prediction theory when quadratic indices are used for the strategic optimisation of breeding schemes where the risk, summarised by ΔF, is constrained. Whilst breeding programmes under BLUP selection can potentially be operated using dynamic selection algorithms, up to date, there is not available framework for the optimal design of breeding schemes under quadratic optimisation.

1.4. Objectives

This General Introduction has highlighted that quadratic indices constitute a general approach for the strategic optimisation of livestock breeding programmes with restricted ΔF as a general measure of risk. Whilst dynamic selection algorithms have been developed for a variety of scenarios (e.g., discrete and overlapping generations scenarios), and have proven to be optimal independently of the genetic model used (i.e., polygenic or mixed inheritance) the following areas have not been fully or at all developed:

i. Although simulation studies have indicated that quadratic indices implemented as dynamic selection tools provide the maximum rate of gain at

a pre-defined rate of inbreeding for a given set of candidates with known relationships and estimated breeding values, the implementation in real livestock breeding programmes is practically not existent. Therefore, there is currently no assessment made on the potential and achievable benefits of quadratic indices when compared to current selection methods (e.g., truncation selection) in ongoing breeding schemes. Also, all previous work has dealt with single trait scenarios although in practice several traits usually contribute to the total aggregate breeding value of a particular individual.

ii. Deterministic predictions of the rate of genetic gain under constrained inbreeding are not available. Therefore, the question of how much rate of gain can be obtained for a given level rate of inbreeding and a set of design variables (e.g., scheme size and trait h^2) cannot be addressed *a priori*, before the breeding programme is started. In consequence, the only way of currently quantifying the benefits from quadratic optimisation over traditional truncation selection is by means of stochastic simulations. Hence, available expectations may not be comprehensive, as they are limited to a rather small number of scenarios.

Therefore this thesis has been designed to target the following objectives:

1. To implement quadratic optimisation and the available dynamic selection algorithms for breeding programmes in existing commercial livestock breeding populations. Applications will be focused on two UK livestock populations of beef cattle (Aberdeen Angus) and sheep (Meatlinc). The following research outcomes were anticipated:

    - A realistic assessment of the expected benefits from quadratic optimisation
    - Identification of problems in the implementation of dynamic tools in large-scale situations (e.g., thousands of candidates rather than hundreds)

- Operational software for the day-to-day operation of commercial livestock breeding programmes

2. To develop a deterministic prediction framework for predicting rates of genetic gain with constrained inbreeding under quadratic indices for breeding schemes under BLUP evaluation of breeding values. The following research outcomes were anticipated:

    - A self-contained prediction of genetic gain for a known rate of inbreeding after specification of the scheme size and trait h^2
    - A framework for evaluating the benefits of quadratic optimisation over truncation selection under extensive combinations of design parameters and risk policies
    - A software with at least the potential to be developed for commercial applications

3. To extend the current evaluation of benefits of quadratic indices in mixed inheritance models from single trait scenarios to multiple trait scenarios. Stochastic simulations were used to study gene assisted selection (GAS) when the dynamic selection tool with constrained inbreeding is applied on an index that includes two negatively correlated traits and an identified QTL affects one of the traits in the index. The following research outcomes were anticipated:

    - Evaluation of the short- and long-term benefits from quadratic optimisation for the overall breeding objective and for each of the traits individually
    - Understanding of the interaction between quadratic optimisation and the use of gene information in schemes under BLUP evaluation of breeding values where multi-trait breeding goals are used.

2. CHAPTER TWO

Expected Increases in Genetic Merit from using Optimised Contributions in two Livestock Populations of Beef Cattle and Sheep

2.1. Introduction

Best Linear Unbiased Prediction (BLUP) has become the standard method for genetic evaluation in breeding programmes of beef cattle and sheep livestock populations. Although selection exclusively based on BLUP estimated breeding values (EBV) allow accurate selection and increased genetic gains, it can also lead to increased rates of inbreeding in comparison with less accurate methods (e.g., Quinton *et al.*, 1992).

While inbreeding is unavoidable in closed selection programmes, increases in inbreeding need to be restricted to alleviate long-term negative effects (Lamberson and Thomas, 1984; Burrow, 1993). Woolliams *et al.* (2002) have described the rate of inbreeding as a measure of risk from the perspective of the breeding programme justifying its management with arguments that go beyond avoiding inbreeding depression and loss of genetic variation in the selected trait.

Dynamic tools for maximising genetic progress whilst constraining the rate of inbreeding to a pre-defined value are now available (Meuwissen, 1997; Grundy *et al.*, 1998; Meuwissen and Sonesson, 1998; Grundy *et al.*, 2000). These tools optimise the number of parents and their contributions to subsequent generations for maximising gain for a fixed rate of inbreeding. Simulation studies showed improvements in genetic gain greater than 20% over BLUP truncation selection at the same rate of inbreeding (Meuwissen, 1997; Grundy *et al.*, 1998). However, the expected benefits from optimised selection in real livestock populations remain unknown.

The main objective of this study was to evaluate the potential extra gains to be obtained by dynamic optimisation algorithms in two livestock populations of sheep (Meatlinc) and beef cattle (Aberdeen Angus). This was accompanied by a description of the population structure, rates of genetic gain and inbreeding, and mating systems before and after the introduction of BLUP evaluations. Relationships between

contributions of ancestors of the current population and their EBV were also evaluated.

2.2. Materials and methods

The Meatlinc (ML) and Aberdeen Angus (AA) breeds were chosen for this study because they have maintained effective improvement programmes and have achieved high genetic gains in the UK when compared with other breeds of sheep and beef cattle, respectively (Simm, 1998; MLC, 1999). Also, in both populations, concerns regarding increasing levels of inbreeding and its potential consequences have arisen (G. Nieuwhof, personal communication).

2.3. Data

The AA is a traditional British beef breed, with a recorded pedigree extending over 50 years. The ML is a synthetic terminal sire breed of sheep created in the UK in the early 1960's. In contrast with AA, the recorded ML pedigree is relatively small, covering only 24 years. Pedigree data and index scores for both populations were provided by the Meat and Livestock Commission (MLC, UK). The two indices provided were the BLUP-EBV for the breeding goals of UK terminal sire breeds of beef cattle and sheep (*'beef value'* and the *'lean index'*, respectively). The *'beef value'* includes carcass weight, carcass conformation score and carcass fat score (Amer *et al.*, 1998; Simm, 1998) while the *'lean index'* includes carcass lean weight and carcass fat weight (Simm and Dingwall, 1989).

The AA pedigree included a total of 119,953 animals (57,431 males and 62,522 females) born from 1948 to 2000. A total of 45,472 parents (6,686 sires and 38,786 dams) were identified. The ML pedigree included a total of 12,391 animals (5,661 males and 6,730 females) born from 1974 to 2000. A total of 3,742 parents (329 rams and 3,413 ewes) were identified. Parents with unknown genealogies were considered as 'base parents'. This group represented 28.9% of the total number of

parents (2,443 sires and 10,704 dams) in AA and 7.4% of the total number of parents (35 rams and 243 ewes) in ML.

Because multi-trait BLUP evaluations were introduced for both populations in 1991, the analyses performed were each applied to three periods of data. The three periods included an overall period covering all years with available information, and two periods of approximately equal length defined pre-BLUP and post-BLUP introduction.

2.3.1. Generation intervals

The generation interval for each breed was computed as the average age of parents at the birth of their offspring. It was calculated for each year of birth and then averaged over years for all parents (L), for sires (L_m) and for dams (L_f).

2.3.2. Rates of genetic progress and inbreeding

Average index scores and inbreeding coefficients of individuals born at each year were calculated. The inbreeding coefficients (F) were obtained from the additive relationship matrix that was computed using the algorithm of Meuwissen and Luo (1992). The rate of genetic gain (ΔG) and the rate of inbreeding (ΔF) were computed as the linear regression of the average index score and average F on the year of birth, respectively.

Both ΔG and ΔF were analysed for the three periods in both populations. For ML, the periods were: i) overall period from 1974 to 2000, ii) pre-BLUP period from 1983 to 1991; and iii) post-BLUP period from 1992 to 2000. For AA, the corresponding periods were: i) overall period from 1948 to 1999; ii) pre-BLUP period from 1983 to 1991; and iii) post-BLUP period from 1992 to 1999.

2.3.3. Long-term genetic contributions

The impact of different cohorts of ancestors on genetic gain was investigated by studying the relationship between their long-term genetic contributions and index scores. The long-term contribution (r) of an ancestor is defined as the proportion of genes it contributes in the long-term to the population (Wray and Thompson, 1990). Over many generations, in a population thoroughly mixed, the r of an ancestor will converge to the same value for all its descendants but will differ among ancestors (Woolliams *et al.*, 1999). Long-term contributions were computed following the approach used by Woolliams and Mäntysaari (1995). To compute r, a generation of ancestors and a generation of descendants were defined according to average generation intervals previously calculated. Thus, the ancestral and descendant generations were defined by using L. This definition ensures that r summed over all ancestors over a period of L years equals unity (Bijma and Woolliams, 1999). Convergence of contributions was assumed if the variance of contributions of ancestors across descendants was lower than 1.0×10^{-4}. For ML (where L was about 2 years), contributions were calculated for two generations of ancestors: i) the cohorts born between 1983 and 1984 and ii) the cohorts born between 1991 and 1992 (i.e., the first generation after the introduction of BLUP evaluation). For both groups of ancestors, descendants were the cohorts born between 1999 and 2000. For AA, ancestors were the cohorts born between 1976 and 1979 (L was about 4 years from 1971 to 1988), and the descendants were the cohorts born between 1995 and 1999 (L was about 5 years from 1988 to 2000). The regression of the long-term genetic contribution of ancestors on their index scores was calculated for each cohort of ancestors.

2.3.4. Mating systems

The degree of non-random mating in both populations was investigated using two approaches. Firstly, the regression of male index score on female index score was estimated together with their correlation. Secondly, for each year cohort, the

deviation (a) of the observed inbreeding coefficient (F_{obs}) from the inbreeding level that would have been expected (F_{exp}) under complete random mating (i.e., deviations from Hardy-Weinberg equilibrium) was obtained. The coefficient a was calculated as $(F_{obs}\text{-}F_{exp})/(1\text{-}F_{exp})$ (Falconer and Mackay, 1996). The F_{exp} of animals born in a certain year was computed as the average co-ancestry of the selected animals born one generation before (i.e., L years before). Although the value of a for a population under complete random mating is zero, some small deviations of the order of –(1/8M + 1/8F) where M and F are the number of males and females, respectively, are expected when the population is divided in sexes (Roberston, 1965). Thus, to assess the degree of deviation from random mating, the observed a coefficients were compared to Robertson's deviation rather than to zero.

2.3.5. *Optimising genetic contributions for maximising genetic gain*

The potential extra genetic gains expected from using selection tools based upon the algorithm described by Meuwissen (1997) were investigated. The algorithm was used to obtain the number of individuals to be selected and the number of offspring each of them should contribute to the next cohort, to achieve the maximum ΔG while constraining ΔF to a specific value. Different restrictions on ΔF were considered. The algorithm maximised the following objective function (Meuwissen, 1997):

$$\mathrm{H}_t = \mathbf{c}_t^{\mathrm{T}}\mathbf{g}_t - \lambda_0(\mathbf{c}_t^{\mathrm{T}}\mathbf{A}_t\mathbf{c}_t - C_t) - [\mathbf{c}_t^{\mathrm{T}}\mathbf{Q} - (1/2)\mathbf{1}^{\mathrm{T}}]\boldsymbol{\lambda}$$

where $\mathbf{c}_t$ is the solution vector of mating proportions (c) of candidates at generation t, $\mathbf{g}_t$ is the vector of EBV of selection candidates, $\mathbf{A}_t$ is the numerator relationship matrix for selection candidates, $\mathbf{Q}$ is a known incidence matrix for the sex of the candidates, $\mathbf{1}^{\mathrm{T}}$ equals [1 1], and λ_0 and $\boldsymbol{\lambda}$ are Lagrangian multipliers. The restriction on the inbreeding rate was achieved each generation by setting $C_t = 2[\Delta F + (1-\Delta F)F_t]$, where F_t is the average inbreeding coefficient of selection candidates. The third term in the objective function ensures that male and female

parents contribute with a half of the gene pool each. Selected candidates are those with $c > 0$ and will contribute to the next generation according to their c value.

The optimisation described above does not take into account any constraint upon the maximum contribution a particular candidate may have that may arise from reproductive limitations. This might not be a problem in males since AI techniques are often widespread in livestock populations. However, it can be unrealistic for female candidates for which high reproductive rates are less feasible, particularly in beef cattle and sheep populations. In order to obtain more realistic results, another set of optimisations was run with an additional constraint on the female contributions. In this case all females were selected by setting their contributions to a pre-defined value (i.e., $1/2n_f$, where n_f is the number of female candidates). This implies that all female candidates are selected and only male mating proportions are optimised. The objective function was modified following Meuwissen (1997, Appendix):

$$H_t = \mathbf{c}_{1_t}^{\mathrm{T}}\mathbf{g}_{1_t} - \lambda_0(\mathbf{c}_{1_t}^{\mathrm{T}}\mathbf{A}_{11_t}\mathbf{c}_{1_t} + 2\mathbf{c}_{1_t}^{\mathrm{T}}\mathbf{A}_{12_t}\mathbf{c}_{2_t} - K_t) - (\mathbf{c}_{1_t}^{\mathrm{T}}\mathbf{Q}_1 - \mathbf{s}^{\mathrm{T}})\lambda$$

where $\mathbf{c}_{1_t}$ is the solution vector of mating proportions of male candidates at generation t, $\mathbf{g}_{1_t}$ is the vector of EBV of male candidates, $\mathbf{A}_{11_t}$ and $\mathbf{A}_{12_t}$ are sub-matrices of $\mathbf{A}$ including only male, and male by female candidates, respectively, $\mathbf{c}_{2_t}$ is the known vector of female mating proportions, K_t is $2C - \mathbf{c}_{2_t}^{\mathrm{T}}\mathbf{A}_{22_t}\mathbf{c}_{2_t}$, $\mathbf{A}_{22_t}$ is the relatedness matrix for female candidates, $\mathbf{s}^{\mathrm{T}}$ is a vector with constant values [0 ½] and $\mathbf{Q}_1$ is a known incidence matrix for males analogous to $\mathbf{Q}$ in the unconstrained case. Software was developed in Fortran 90 to solve the objective functions described above.

Potential benefits from using optimised contributions were estimated by comparing the expected index gains obtained by using the selection algorithm after mimicking selection in 1999 to: i) the actual observed ΔG in 2000; and ii) the expected ΔG in 2000 under truncation selection (i.e., equal contributions) at the observed ΔF in the population being evaluated. The expected ΔG from truncation selection was

calculated by allocating a fixed mating proportion to female candidates (i.e., equivalent to one mating) and by selecting the number of male candidates that gave the observed ΔF. This latter comparison allows evaluating the expected benefits from optimising contributions independently of the benefits of selecting solely on the index. The fact that in practice selection intensity might be lower than that achievable if selection decisions include other than exclusively BLUP-EBV criteria (e.g., Lewis and Simm, 2000) is not accounted for in comparison i).

Candidates for the selection algorithms were defined by using both L_m and L_f. Therefore, for ML, candidates were those males born in 1999 (L_m = 1.0 years) and those females born from 1996 to 1998 inclusive (L_f = 3.0 years). The total number of candidates was 1,841. For AA, candidates were those males and females born from 1992 to 1998 inclusive ($L_m = L_f$ = 5.0 years) and this gave a total number of candidates of 55,553. However, in order to reduce computing requirements a pre-selection of candidates was performed by imposing a minimum index score. For ML, a total of 1,297 candidates (395 males and 902 females) with index score equal or greater than 179.0 were included. For AA a total of 6,429 candidates (3,321 males and 3,108 females) with index score equal to or greater than 21.0 were included. When only male mating proportions were optimised computer requirements were higher and in this case, only 417 male candidates (those with index score equal or higher than 30.0) were included in the AA optimisation. The index scores were those obtained from the MLC genetic evaluation in 2000, and ΔF was constrained to a range of values including the observed inbreeding rate per generation in each breed.

2.4. Results

The pre- and post-BLUP periods are indicated in the figures presenting results on population structure (Figure 2.1), generation intervals (Figure 2.2), rates of genetic gain and inbreeding (Figure 2.3), and mating systems (Figure 2.5). The total number of years analysed in each case depended on the available information, but the pre-

and post-BLUP periods are indicated according with the definition given in the Methods section.

2.4.1. Population structure

Table 2.1 shows descriptive statistics summarising the population structure for both populations. The number of ML rams and ewes, and the ewe to ram ratio (d) per year are shown in Figure 2.1a for the period 1974 to 2000. A large increase in the number of ewes per ram was observed from 1974 (d = 4.5) to 2000 (d = 24.4), although the ratio remained more or less constant for the period after the introduction of BLUP. The breed showed an important expansion through a steady increase in the number of ewes from 1981, from about 50 to about 700 in 2000. The increase in the number of rams was however moderate from about 5 in 1974 to about 30 in 2000.

For AA, a steady increase in the number of dams per sire was observed from 1969 (d = 2.8) to 1999 (d = 10) (Figure 2.1b). The number of breeding animals increased substantially from 1984 particularly the number of dams, which showed a fivefold increase. In contrast with the ML case, this might be due to an increase in the breed membership to the recording services rather than to a genuine breed expansion.

The average number of offspring per male across years in AA (18.4) was very close to the upper bound of the 25% to 75% inter-quartile range (2 to 19; Table 2.1), indicating a much more skewed distribution than for ML where the average (37.6) falls near the mid-point of the range (27 to 43; Table 2.1).

2.4.2. Generation intervals

Figure 2.2 shows the average generation interval over years for males (L_m) and females (L_f) for ML and AA. In ML, L_m was calculated from 1983 onwards because ram dates of birth were not available before this year. An important increase in L_f over years was observed in the period 1976 to 1983. This increase is related to the

period of establishment of this synthetic breed in which females had to be kept in the flock for more time. From 1984 onwards, L_f remained unchanged and the average was 3.2 years. In this population, L_m remained unchanged around a value of 1.0 years over the pre-BLUP period but has slightly increased over the post-BLUP period up to 1.4 years.

In AA both L_m and L_f increased at similar rates (0.11 ± 0.02 and 0.16 ± 0.01 per year, respectively) during the period 1976 to 1987 (Figure 2.2b). The average L_m and L_f in this period were 4.6 and 4.7 years, respectively. Over the last 12 analysed years (1988 to 2000) L_m and L_f averaged 5.2 and 5.7 years, respectively, although since 1994 the generation intervals started to diverge. By year 2000, L_f was around one year larger than L_m. There was no evidence to link this increase in L_m and the use of BLUP-EBV.

2.4.3. Rates of genetic progress

Figure 2.3 shows the average index values per year of birth for ML and for AA for the periods 1982 to 2000 and 1970 to 1999, respectively. Results indicate that the introduction of BLUP evaluations led to a sustained increase in the rate of genetic gain in ML from 1994. The difference between ΔG in the pre- and post-BLUP periods was statistically significant. For this breed, ΔG was 5.5 ± 1.0 ($P < 0.01$) index units per year in the pre-BLUP period, and 16.5 ± 0.6 ($P < 0.01$) index units per year in the post-BLUP period. On the other hand, the ΔG for AA before and after the BLUP introduction were not significantly different. The pre-BLUP and post-BLUP rates of gain were 0.55 ± 0.04 ($P < 0.01$) and 0.46 ± 0.05 ($P < 0.01$) index units per year, respectively.

2.4.4. Rates of inbreeding and long-term contributions

The average inbreeding coefficient (F) in the ML population in the year 2000 was 6.3% (Figure 2.3a). The ΔF per year for the period 1982 to 2000 was 0.19% ($P <$

0.001). The difference between ΔF in the pre- and post-BLUP periods, 0.21% ± 1.31% and 0.23% ± 0.05%, respectively, was not significant. Nevertheless, the pre-BLUP estimation of ΔF should be taken with caution as F fluctuated considerably in this period. Considering that the generation interval of the population in the post-BLUP was about 2.3 years (Figure 2.2a), the ΔF per generation in this period was 0.53%. This is equivalent to an effective size of the population (N_e) of 95 animals (i.e., $N_e = 1/2\Delta F$). On the other hand, the ΔF increased in the last generation up to about 1.0% (i.e., $N_e = 50$).

The average F in the AA population in 1999 was about 0.97% (Figure 2.3b). For the period 1974 to 1999, ΔF was 0.04% per year ($P < 0.001$). As with ML, the rates of inbreeding for the pre and post-BLUP periods were similar (0.02% ± 0.008% and 0.03% ± 0.008%, respectively). Considering that the generation interval was about 5 years (Figure 2.2b), ΔF per generation in the post-BLUP period was approximately 0.15% (i.e., N_e = 333). The ΔF in the last generation (i.e., from 1994 to 1999, inclusive) was about 0.20%.

For ML, 203 ancestors born between 1983 and 1984 (21 males and 182 females) were identified for computation of their long-term contributions to the 2,094 descendants born between 1999 and 2000. The relationship between long-term genetic contributions of these ancestors and their index values is shown in Figure 2.4a. The regression coefficients of contributions on index scores were not significant for all ancestors (4.1×10^{-5}; $P = 0.15$), nor for ram ancestors (-1.6×10^{-4}; $P = 0.13$), but was significant for ewes (6.8×10^{-5}; $P < 0.01$). For the analysis of contributions for the first generation after BLUP evaluation, long-term genetic contributions for the 1,337 ancestors born from 1991 to 1992 (643 males and 694 females) to descendants born between 1999 and 2000 were computed. In this case, the regressions of contributions on index scores for this set of ancestors were significant for both males (3.6×10^{-5}; $P < 0.01$) and females (2.4×10^{-5}; $P < 0.01$).

For AA, a total of 5,861 ancestors (2,686 males and 3,175 females born between 1976 and 1979) and 48,248 descendants (born between 1995 to 1999) were

identified. The regressions of contributions on index score were 3.3×10^{-6} ($P < 0.001$) for all ancestors, 5.7×10^{-6} ($P < 0.001$) for male and 1.3×10^{-6} ($P < 0.03$) for female ancestors. The analysis of contributions of male ancestors having long-term contributions greater than zero only, gave a higher regression coefficient (9.3×10^{-5}; $P = 0.01$) than that obtained with all male ancestors. No additional long-term contribution analysis (i.e., for the first generation after BLUP) was carried out due to the long L, which would have implied non-converged contributions.

2.4.5. *Mating systems*

There was evidence of positive assortative matings with respect to index scores in both populations. The regression of rams index value on ewes index value and the correlation between them was 1.02 ($P < 0.001$) and 0.70, respectively for ML and 0.41 ($P < 0.001$) and 0.33, respectively for AA. The higher correlation observed in ML suggests a more co-ordinated mating policy, possibly enabled by the smaller scale of the breeding programme.

Evidence of departures from random mating with respect to the pedigree is suggested for ML (Figure 2.5a). For this population, the average expected a values for random mating according to Robertson's deviations before and after the introduction of BLUP were –1.25% and –0.57%, respectively. Avoidance of matings among relatives seems to have been practised in the pre-BLUP period (e.g., average observed a of –2.40%), in particular between 1982 and 1987 (e.g., average a of –3.00%), possibly related to the period of formation of this synthetic breed. On the other hand, during the post-BLUP period, the average a (e.g., -1.2%) was much closer to Robertson's deviation, giving no clear indication of departures from random mating. This suggests a lower level of population subdivision compared to the pre-BLUP period that could be related to the creation of genetic links among flocks to enable across BLUP evaluations through the use of common rams. The overall range of deviations of observed a with respect to random mating expectation observed in ML was from 0.0% to –5.0%. This range is of the same magnitude of those observed

by Maroo (1999) and Nomura *et al.* (2001) when describing the population structures of the Exmoor pony and of the Japanese Black cattle populations, respectively.

For AA, although a steady increase in the a coefficient throughout the period analysed was observed (Figure 2.5b) there was no clear indication from deviations from random mating with respect to the pedigree. The overall scale of a for this population was about ten times lower than for ML (see also Figure 2.5a). The average pre-BLUP and post-BLUP observed a coefficients were 0.23% and –0.55%, respectively. Given the size of the population, Robertson's expected deviations were negligible (e.g., ranging from -0.02% to -0.03%).

2.4.6. *Expected increases in genetic merit from applying optimised selection*

Table 2.2 shows the optimisation results for both populations. Three levels of ΔF constraints (0.5%, 1.0% and 2.0%) were considered for ML, and four levels (0.2%, 0.5%, 1.0% and 2.0%) were considered for AA. The observed ΔF at the last generation were 1.0% for ML and 0.2% for AA.

2.4.6.1. Optimisation of contributions in both sexes

The optimisation of contributions of both male and female candidates led to substantial increases in predicted average index score ($\text{Index}_{\text{P-2000}}$) and index gain ($\Delta G_{\text{P-2000}}$) in 2000 in both populations (Table 2.2). The observed index gain ($\Delta G_{\text{O-2000}}$) from 1999 to 2000 was 16.1 index units for ML and 3.6 index units for AA (see Table 2.2). When ΔF was restricted to the ΔF observed in the last generation, the $\Delta G_{\text{P-2000}}$ were 70.2 index units (i.e., 4.4-fold over $\Delta G_{\text{O-2000}}$) for ML ($\Delta F = 1.0\%$) and 21.1 index units (i.e., around sixfold over $\Delta G_{\text{O-2000}}$) for AA ($\Delta F = 0.2\%$). Further relaxation of the ΔF constraint led to higher increments in index gain. For instance, at the most relaxed constraint ($\Delta F = 2.0\%$) the expected increments over the observed ΔG were 65.6 index units for ML and 22.1 index units for AA. However,

the relaxation in the restriction on ΔF contributed only to relatively small increases in ΔG in comparison to the increases observed by optimising contributions. It should be noted that for ML an increment in ΔG of 46.5 index units was expected even if ΔF was restricted to a value as low as 0.5%, which was the ΔF observed over the post-BLUP period.

At the tightest constraint in ΔF, the number of selected candidates was 80 (i.e., 31 males and 49 females) for ML and 149 (i.e., 68 males and 81 females) for AA. As expected, as the ΔF restriction was less severe the number of selected candidates decreased. For the most relaxed constraint (ΔF = 2.0%) it dropped down to 55 (i.e., 18 males and 37 females) for ML and to 36 (i.e., 17 males and 19 females) for AA.

2.4.6.2. Optimisation of male contributions when all females are selected.

For a more realistic scenario for typical production systems where female contributions are restricted, the algorithm still achieved significant predicted increases in index score gains (Table 2.2). At the observed ΔF in the last generation, $\Delta G_{P\text{-}2000}$ was 32.2 index units (i.e., twofold over $\Delta G_{O\text{-}2000}$) for ML (ΔF = 1.0%) and 15.2 index units (i.e., around fourfold over $\Delta G_{O\text{-}2000}$) for AA (ΔF = 0.2%). As when contributions of both sexes were optimised, the relaxation of the constraint allowed for even higher predicted increases. For the most relaxed constraint (ΔF = 2.0%), predicted increases in index gain were 32.6 index units for ML and 14.8 index units for AA.

When ΔF was restricted to a lower value than that observed in the last generation (i.e., 0.5%), $\Delta G_{P\text{-}2000}$ was 8.6 index units lower than the observed index gain in 2000 as it implied a very tight constraint (Table 2.2).

Because female contributions were fixed, the expected relative gains over the observed average index scores arose only from the management of the male selection intensity. The number of selected rams in ML decreased by relaxing the constraint on

ΔF from 58 (ΔF = 0.5%) to 19 (ΔF = 2.0%), while for AA the number of selected bulls decreased from 67 (ΔF = 0.2%) to 11 (ΔF = 2.0%) (Table 2.2).

The predicted benefits over the observed ΔG might be overestimated as at the current ΔF, breeders may be able to achieve higher selection intensities if selection were based solely on index values. The rates of gain under truncation selection (i.e., with equal mating proportions) that gave the observed ΔF were 27.6 index units for ML and 11.7 units for AA (Table 2.2). Thus, benefits of optimal selection when only male contributions were optimised over truncation selection (based exclusively on EBV) at the observed ΔF were 17% for ML and 30% for AA. This suggests that the above twofold and fourfold expected benefits in ΔG for ML and AA, respectively, may be over predictions.

2.4.7. *Relationship between optimised mating proportions and index scores*

To achieve the restriction on ΔF, the more severe the ΔF constraint was, the higher was the number of selected candidates and the more alike were their optimised mating proportions. This behaviour across ΔF constraints can be seen in Figure 2.6 for AA where selection was mimicked at year 1999. The same pattern of behaviour was observed for ML.

As the ΔF constraint was relaxed, the variance of optimised mating proportions among selected candidates increased from 4.3×10^{-5} (ΔF = 0.2%) to 4.9×10^{-4} (ΔF = 2.0%), whereas the variance of the index score among selected candidates decreased from 23.0 (ΔF = 0.2%) to 13.6 (ΔF = 2.0%). The highest optimal mating proportion was assigned to the individual with the highest index score (5.20 units) and ranged from 0.042 (ΔF = 0.2%) to 0.102 (ΔF = 2.0%).

2.5. Discussion

This work has shown that the population structure of the two pedigree breeds analysed has changed across years and suggests that, in at least one of the populations (i.e., the ML), the introduction of BLUP has led to sustained additional genetic gains. Concomitant increases in the average coefficient of inbreeding have been observed, although there was no evidence that ΔF was increased by the introduction of BLUP. The application of dynamic selection tools for maximising genetic gain while constraining rates of inbreeding to target levels would have led to important benefits in ΔG compared to what has been observed and over the expected gain under truncation selection based exclusively on index scores. This demonstrates the scope for this type of optimisation tool in livestock breeding programmes.

2.5.1. The impact of artificial insemination (AI)

The impact of AI on both populations can be clearly seen in the increase in mating ratios of the breeding males across years. Nowadays, this technique is a standard reproductive technique in beef cattle, however, in general it is less widespread in sheep populations where the AI procedures are much more complex and success rates are typically much lower. Nevertheless, AI techniques in the UK have been promoted as a key element for the establishment of Sire Reference Schemes (SRS). The SRS enabled BLUP evaluations across flocks, and thereby increased the potential benefit from the use of BLUP (Simm, 1998). The widespread use of some sires has not only led to an increase in d but also in the variance of the number of offspring per male (results not shown).

It might be anticipated that increases in the number of offspring per male and its variance would have led to increases in ΔF. However, ΔF has remained relatively steady, particularly in the AA. This has been due to the expansion of the recorded breed numbers, and in particular to an increase in the number of bulls used per year in the population over the period studied. This simple step has reduced the

proportional contributions of individual males to the gene pool and so limited the expected increase in ΔF.

2.5.2. The impact of BLUP

The introduction of BLUP evaluations seems to have led to an increase in ΔG for the ML, but this response was not observed in the AA. One reason for this difference is the structure of the populations: whereas the ML consists of four closely co-operating flocks, with selection policies closely defined by the selection index (H. Fell, personal communication), the AA has a looser breeding pyramid with about 200 herds, where policies of individual breeders might not be so closely determined by the Society alone. Also, in beef cattle, there is likely to be a much higher use of older 'proven' males via AI than in sheep. The ML may be better placed to utilise the more accurate information arising from the BLUP evaluation, and in combination with AI, the better information across flocks produced by the SRS. A further example of the impact of BLUP in the ML is the evidence of increases in L_m during the post-BLUP period where ram usage was extended for a longer period because of the better comparison across age groups made possible by BLUP.

The exclusive use of BLUP-EBV as a ranking tool for truncation selection would be expected to lead to an increase in ΔF (Quinton *et al.* 1992), due an increased co-selection of relatives. Nevertheless, this phenomenon was not observed in either breed during the post-BLUP period, which may be particularly surprising in the ML where closer attention was paid to the index evaluations. However, the increase in breeding males used per year in both populations has proved effective in managing ΔF over the short-term. The changes in L_m observed post-BLUP for ML would have led to a reduction in ΔF per year, but this benefit would be offset by the larger lifetime genetic contributions arising from rams kept more than one year.

The estimated effective population sizes in the post-BLUP period are 95 for ML and 333 for AA (i.e., ΔF of 0.53% and 0.15% per generation, respectively). These values

are above the minimum reference value of 40 animals of Goddard and Smith (1990) for maximising net genetic response for total economic merit in dairy cattle, and fall within the critical range of 30 to 250 animals of Meuwissen and Woolliams (1994) for balancing decreases in fitness due to inbreeding and increases in fitness due to natural selection. Nevertheless, there was a substantial decrease in effective size in ML in the last generation to a value comparable to the minimum effective size of 50 recommended by FAO (1998). Thus, the application of methods for avoiding further future increases in ΔF in this population is advisable.

When ML ancestors born from 1991 to 1992 were analysed, those with higher index values tended to have larger long-term contributions as regressions on index values were positive and significant, compared to a more uniform relationship between ram usage and index scores during the early establishment phase of this synthetic population. This result clearly coincides with the higher genetic gains achieved after the implementation of BLUP-EBV in 1991. A positive association of contributions with index EBV was also observed in AA for the only set of ancestors analysed (i.e., born from 1976 to 1979). For this breed, a comparison of the distribution of contributions before and after the introduction of BLUP is difficult since at most two generations have passed since the introduction and long-term genetic contributions require five or more generations to achieve a reasonable degree of stability. Higher regressions were observed for male ancestors in both populations indicating higher selection intensities applied on males than on females in accordance with expectations (Woolliams *et al.*, 1999).

2.5.3. *Impact of optimised selection*

Although the ML has increased its rate of gain using BLUP evaluations, and simultaneously managed its ΔF by increasing the number of males selected per year (so reducing the selection intensity), further gains are possible by using the selection algorithms. The results showed that these algorithms would benefit both AA and ML over a range of values of ΔF. The most dramatic increases in ΔG were obtained

when selection was allowed in both males and females. However, these gains assume unrealistic reproductive rates for females. Substantial and valuable increases in ΔG were obtained when no selection among females was allowed. At the ΔF in the last generation, the benefits over the observed ΔG were 16.1 index units (i.e., twofold) for ML and 11.6 index units (i.e., threefold) for AA.

The comparisons of the expected (from optimised mating proportions) with observed index scores, implicitly assumed that in practice selection in the two populations has been exclusively based on index values. In practical breeding schemes however, selection decisions are based not only on EBV but also on other factors (e.g., physical and reproductive soundness). This reduces the selection intensity and the maximum genetic merit achievable. For instance, Lewis and Simm (2000) found that losses in selection intensity in sheep SRS can lead to genetic response 0.58 to 0.69 times that obtained when strictly the best animals were selected on BLUP- EBV. Results in ML support this expectation, where the ratio $\Delta G_{O\text{-}2000}/\Delta G_{T\text{-}2000}$ was 0.58 (see Table 2.2). This effect was more important in AA where the ratio $\Delta G_{O\text{-}2000}/\Delta G_{T\text{-}2000}$ was only 0.31. In this case, because of the scale of population, selection decisions might be restricted within herds or groups of breeders. In contrast, the ML SRS is managed as a single selection unit with a single selection policy and tight co-operation among flocks (H. Fell, personal communication). The predicted benefits of 17% for ML and 30% for AA over the expected gain under truncation selection provides an even more realistic evaluation of the benefits of optimal selection when compared to traditional truncation selection based solely on BLUP-EBV. In addition, these expected benefits are in line with the deterministic predictions of the rate of genetic gain at pre-defined rates of inbreeding (see Chapter 4) ranging from 20% to 40% for $\Delta F = 1.0\%$.

Interestingly, ΔG from optimal selection after optimising only male contributions of 15.2 index units (see Table 2.2) at the observed ΔF in AA was similar to the expected ΔG (i.e., 15.7 index units, result not shown) after optimising the contributions of selected bulls in 1999 with observed offspring in 2000 conditional to the observed dam contribution. This indicates that the expected extra index gain from

optimal selection are realistic and would be achievable by only optimising the usage of the current set of bulls selected by the breeders using the available selection index.

In practice, about 30 rams and 700 ewes are used each year in ML, but when contributions of both sexes were optimised, the optimisation algorithm implied selection of 25 males and 43 females on average (Table 2.2). Similarly, in AA the actual number of breeding animals (about 1,000 males and 10,000 females) are much larger than those obtained after applying the optimisation tool on both sexes (43 males and 49 females on average, Table 2.2). As it was pointed out before, these optimum numbers of selected candidates imply very high reproductive rates, particularly on the female side. On the other hand, much more realistic outcomes from the application of the optimisation tool were obtained when only male contributions were optimised conditional on fixed female contributions. At the observed ΔF, the maximum contribution allocated to a male was about 0.046 and 0.048 for ML and AA, respectively. This is equivalent to an expected maximum optimum number of matings per male of 83 (i.e., 2 x 0.046 x 902) for ML and 298 (i.e., 2 x 0.048 x 3,108) for AA. These optimum numbers and the differential usage arising from optimising contributions can be readily achieved through AI, which is currently a standard male reproductive technique. Moreover, they imply a maximum number of offspring per male that is well below the observed upper limit for the number of offspring per male range in each breed (see Table 2.1).

The change in the slope and intercept of the regression coefficient between mating proportions and index scores with different restrictions on ΔF observed here agrees with the general form of the optimal solutions stated by Grundy *et al.* (1998). Basically, as a less severe constraint was imposed, fewer individuals were selected, the usage of the individuals became more unequal, and the slope of the regression was higher (Figure 2.6). Although in the initial cohorts after applying the dynamic selection tools the selective advantage may be the index EBV as suggested by 2.6, once the use of the dynamic selection algorithm is established, the selection advantage is given by the estimated Mendelian sampling term of the index (Woolliams *et al.*, 2002).

Weigel and Lin (2002) applied the algorithm of Meuwissen (1997) in five major US dairy breeds concluding that genetic gain may be sacrificed by imposing constraints on inbreeding. Their conclusion could be somewhat misleading because they did not compare the predicted average genetic merit from the optimisation to that obtained without optimisation at the observed level of inbreeding. Our results clearly indicate that no losses of gain are expected when the inbreeding rate is constrained to the observed value, but additional gains are expected. The only situation where a lower relative gain with respect to the observed gain was obtained, occurred in the extreme case when a constraint equivalent to a lower than the observed ΔF in ML was applied, after fixing all female contributions.

The practical realisation of the optimal contributions and the average expected index scores requires a co-ordinated policy of the use of selected candidates among the different breeding flocks or herds making up the breeding population. This could be a reasonable target in small-scale schemes with co-ordinated breeding policies, but clearly would be much more difficult in large schemes where different objectives might be pursued. Hence, in breeds with large-scale breeding programmes, a more reasonable approach would be to apply the optimisation tool on individual herds or groups of herds with co-ordinated selection policies and objectives.

Two methodological aspects of this optimisation should be addressed. Firstly, the use of EBV obtained in 2000 to mimic selection in 1999, instead of using the EBV obtained in 1999, is not expected to have affected the results obtained as no significant changes in candidates ranking were found when EBV from both evaluations were compared. Secondly, the pre-selection of candidates would not have affected the optimisation outcome. Amongst the group of higher merit candidates those in the bottom half were never selected indicating that any discarded candidate would not have made a significant contribution if it had been included in the optimisation. On the other hand, this allowed a significant reduction of computational requirements.

More flexible constraints may be required in breeding programmes with particular features or breeding structure. These may include setting a maximum contribution per male (i.e., a minimum number of sires), a fixed contribution for a particular set of males, or a desired contribution of a group of females (e.g., in a nucleus). These constraints can be accommodated with the same tool used here (after Appendix in Meuwissen, 1997). For instance, for ML, an additional optimisation with a maximum number of 20 female mates per selected ram (i.e., a minimum of 45 selected rams) was set (not shown). Accordingly, the selection tool found a feasible solution by selecting 46 males, of which 45 were allocated a maximum fixed contribution of 0.01. Moreover, even with this highly restrictive constraint, the expected ΔG at the current ΔF was still about 6.0 index units higher than the observed ΔG.

Although the optimisation approach used here can be realistically applied in practical livestock breeding programmes, evolutionary computation strategies may provide a more flexible framework for setting a greater variety of constraints. Genetic algorithms have been used as optimisation tools in livestock breeding programmes (e.g., Shepherd and Kinghorn, 1998, Meszaros *et al.*, 1999), and they could be extended to explicitly restrict ΔF (e.g., Correnti, 2002).

After allocating optimal mating proportions to the selected candidates, the following step in a breeding programme is to decide a mating policy. Sonesson and Meuwissen (2000) found that the optimisation tool used here combined with mating systems that restricted either mating pairs co-ancestry or offspring co-ancestry, achieved 22% higher response than random mating, in particular for stringent constraints on ΔF. Nevertheless, the extra benefit from the use of non-random matings was reduced as the size of the scheme increased.

Our results refer to practical livestock populations in which the main objective is to achieve the highest genetic gain for a given ΔF. However, the approach is also valid for conservation purposes where the aim may be to minimize ΔF while achieving a pre-defined level of genetic gain (Villanueva *et al.*, 2003). Hence, genetic improvement and conservation can be taken as the extremes of a broader

optimisation problem with particular relative emphasis given to the gain and inbreeding.

This work demonstrates that the application of dynamic optimisation tools allows the management of the rate of inbreeding without any concomitant loss in genetic gain. At the observed rate of inbreeding, substantial benefits were predicted over the expected genetic gains under truncation selection based exclusively on index values and indeed over the observed gains in these two distinct populations of sheep and beef cattle. Breeders have now the opportunity of explicitly managing the risk associated with inbreeding, and to adopt breeding policies according with their risk preferences. The only inputs needed to apply the tool are the estimated breeding values, currently available from genetic evaluations, and an estimate of inbreeding level in the population. The realisation of the benefits from the application of dynamic selection tools requires a co-ordinated policy on the use of selected candidates among the different breeding flocks or herds making up the breeding population.

2.6. Tables

Table 2.1. Summary of females to males ratio (d), number of offspring per male and female parent and generation intervals (L for overall; L_m for males and L_f for females) for Meatlinc and Aberdeen Angus

Parameter	Meatlinc	Aberdeen Angus
[1]d	17.7	6.8
Offspring/male		
25% to 75% range	27 – 43	2 – 19
Average	37.6	18.4
SD	23.7	49.1
Maximum	164	1,093
Offspring/female		
25% to 75% range	2 – 4	2 – 6
Average	3.6	3.2
SD	2.6	1.6
Maximum	21	42
[2]L	2.15	5.09
L_m	1.08	4.95
L_f	3.22	5.24

[1] The average d was calculated from 1974 to 2000 for Meatlinc and from 1969 to 1999 for Aberdeen Angus.

[2] The average L was calculated from 1983 to 2000 for Meatlinc and from 1976 to 2000 for Aberdeen Angus.

Table 2.2. Observed index gain ($\Delta G_{O\text{-}2000}$), predicted average index value ($Index_{P\text{-}2000}$), predicted index gain ($\Delta G_{P\text{-}2000}$) and number of selected candidates[1] after applying optimised selection with different constraints on the rate of inbreeding (ΔF, %) and predicted index gain under truncation selection ($\Delta G_{T\text{-}2000}$) in year 2000 for the Meatlinc and Aberdeen Angus populations

	[2]ΔF			
Item	0.2	0.5	1.0	2.0
Meatlinc				
[3]$\Delta G_{O\text{-}2000}$			16.1	
Both sexes optimised				
$Index_{P\text{-}2000}$		285.9	293.5	305.0
[4]$\Delta G_{P\text{-}2000}$		62.6	70.2	81.7
Selected males		31	26	18
Selected females		49	45	37
Only males optimised				
$Index_{P\text{-}2000}$		230.8	255.5	272.0
$\Delta G_{P\text{-}2000}$		7.5	32.2	48.7
Selected males		58	39	19
Selected females		902	902	902
[5]$\Delta G_{T\text{-}2000}$			27.6	
Aberdeen Angus				
$\Delta G_{O\text{-}2000}$	3.6			
Both sexes optimised				
$Index_{P\text{-}2000}$	39.1	40.4	41.9	43.7
$\Delta G_{P\text{-}2000}$	21.1	22.4	23.9	25.7
Selected males	68	51	36	17
Selected females	81	58	37	19
Only males optimised				

$Index_{P\text{-}2000}$	33.3	34.3	35.4	36.4
$\Delta G_{P\text{-}2000}$	15.2	16.3	17.4	18.4
Selected males	67	41	21	11
Selected females	3,108	3,108	3,108	3,108
$\Delta G_{T\text{-}2000}$	11.7			

[1] Selection candidates included 395 male and 902 female candidates in Meatlinc and 3,321 male and 3,108 female candidates in Aberdeen Angus.

[2] The observed ΔF in the last generation for Aberdeen Angus and Meatlinc was 0.2% and 1.0%, respectively.

[3] The $\Delta G_{O\text{-}2000}$ in each population was calculated from the observed average index values in 1999 and 2000: 223.3 and 239.4 index units, respectively, for Meatlinc, and 18.0 and 21.6 index units, respectively, for Aberdeen Angus.

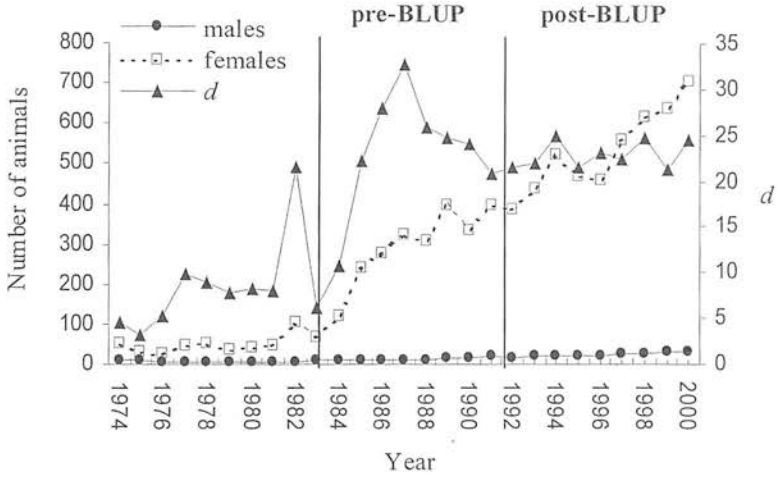
[4] The $\Delta G_{P\text{-}2000}$ in each population was calculated from the corresponding observed average index value in 1999 and the $Index_{P\text{-}2000}$.

[5] The $\Delta G_{T\text{-}2000}$ in each population was calculated from the corresponding observed average index values in 1999 and the average index value calculated with equal contributions of males while fixing female contributions equivalent to one mating. The number of males selected was that giving the observed ΔF in each population: 250.9 index units for Meatlinc and 29.7 index units for Aberdeen Angus.

2.7. Figures

Figure 2.1. Number of male and female parents and ratio dams to sire (*d*) across years for Meatlinc (1974 to 2000) and Aberdeen Angus (1969 to 1999). The pre- and post-BLUP periods are indicated.

a. Meatlinc



b. Aberdeen Angus

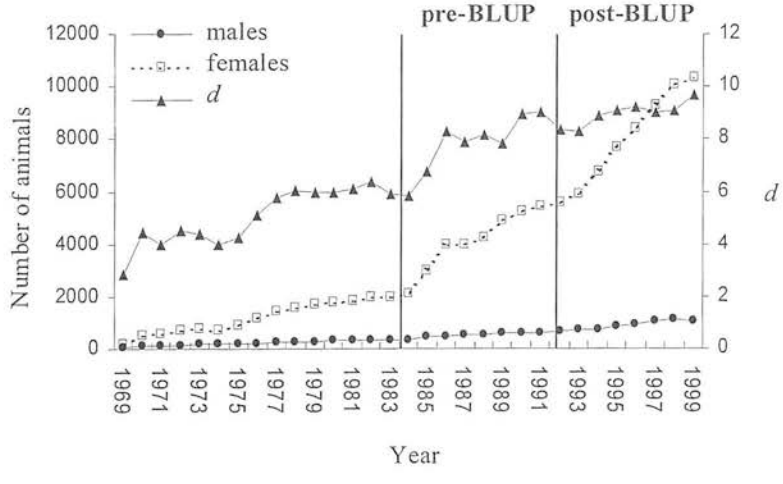
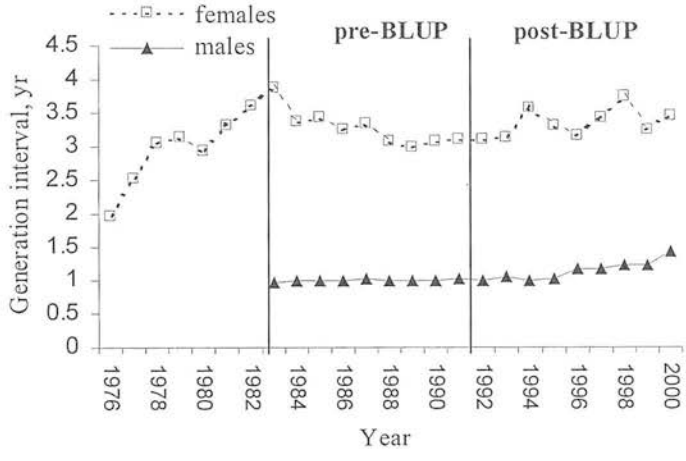


Figure 2.2. Male and female average generation intervals across years for Meatlinc (1976 to 2000) and Aberdeen Angus (1976 to 2000). The pre- and post-BLUP periods are indicated.

a. Meatlinc



b. Aberdeen Angus

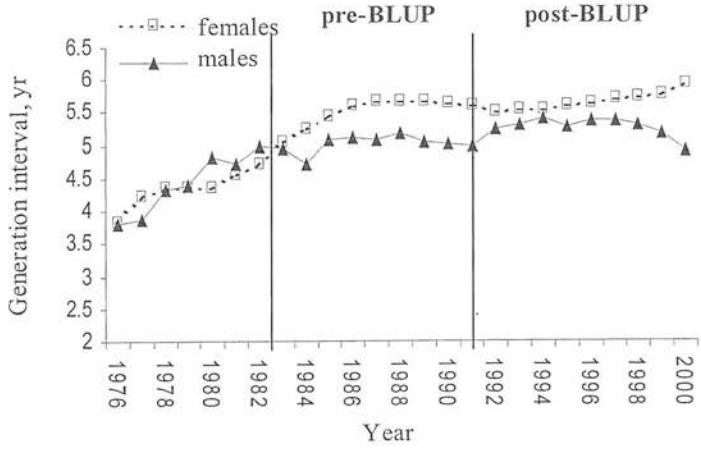
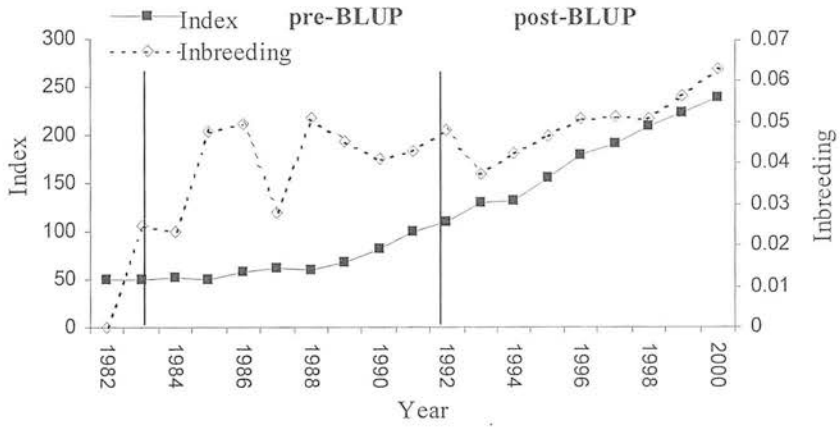


Figure 2.3. Average index score and inbreeding coefficient across years for Meatlinc (1982 to 2000) and Aberdeen Angus (1970 to 1999). The pre- and post-BLUP periods are indicated.

a. Meatlinc



b. Aberdeen Angus

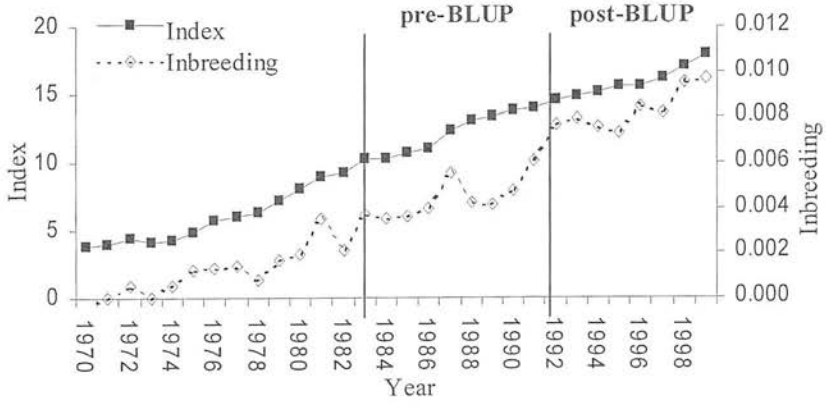
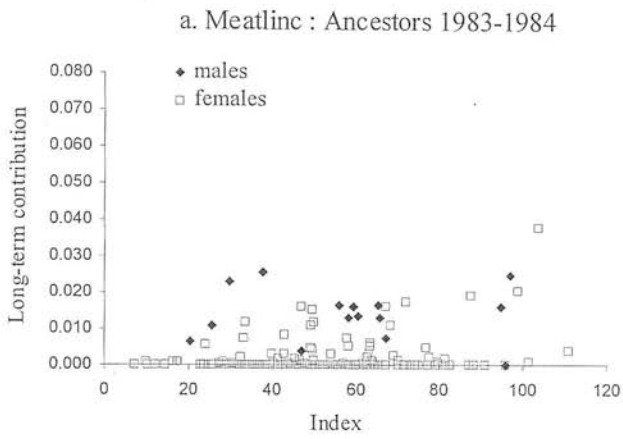
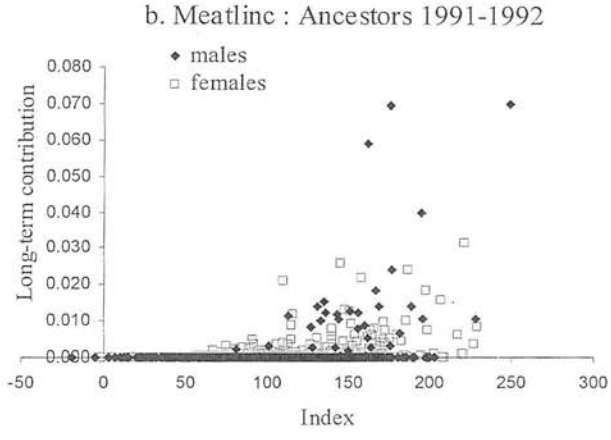


Figure 2.4. Relationship between long-term genetic contributions and index scores of selected male and female ancestors for Meatlinc and Aberdeen Angus.





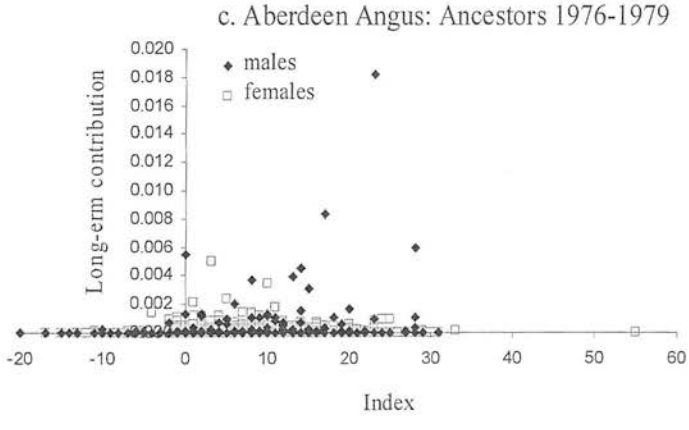
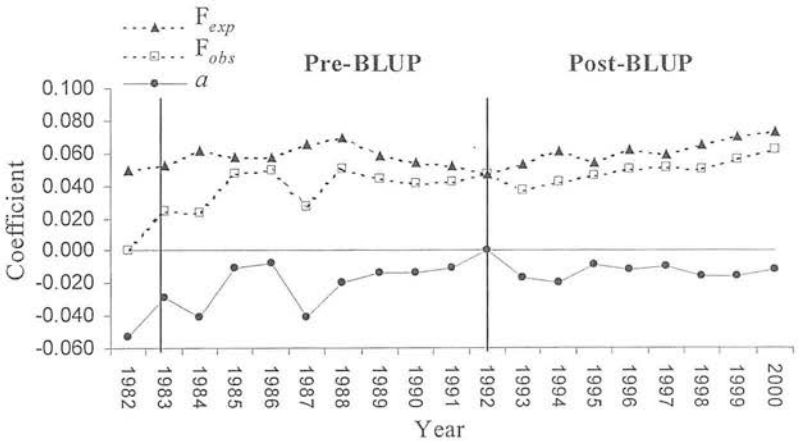


Figure 2.5. Coefficient *a* of deviations from the inbreeding level that would have been expected under complete random mating, mean observed (F_{obs}) and expected (F_{exp}) inbreeding coefficient for Meatlinc (1982 to 2000) and Aberdeen Angus (1975 to 1999). The pre- and post-BLUP periods are indicated.

a. Meatlinc



b. Aberdeen Angus

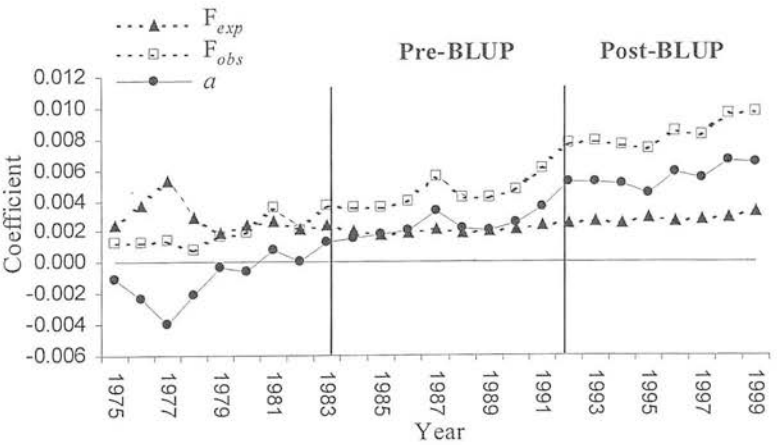
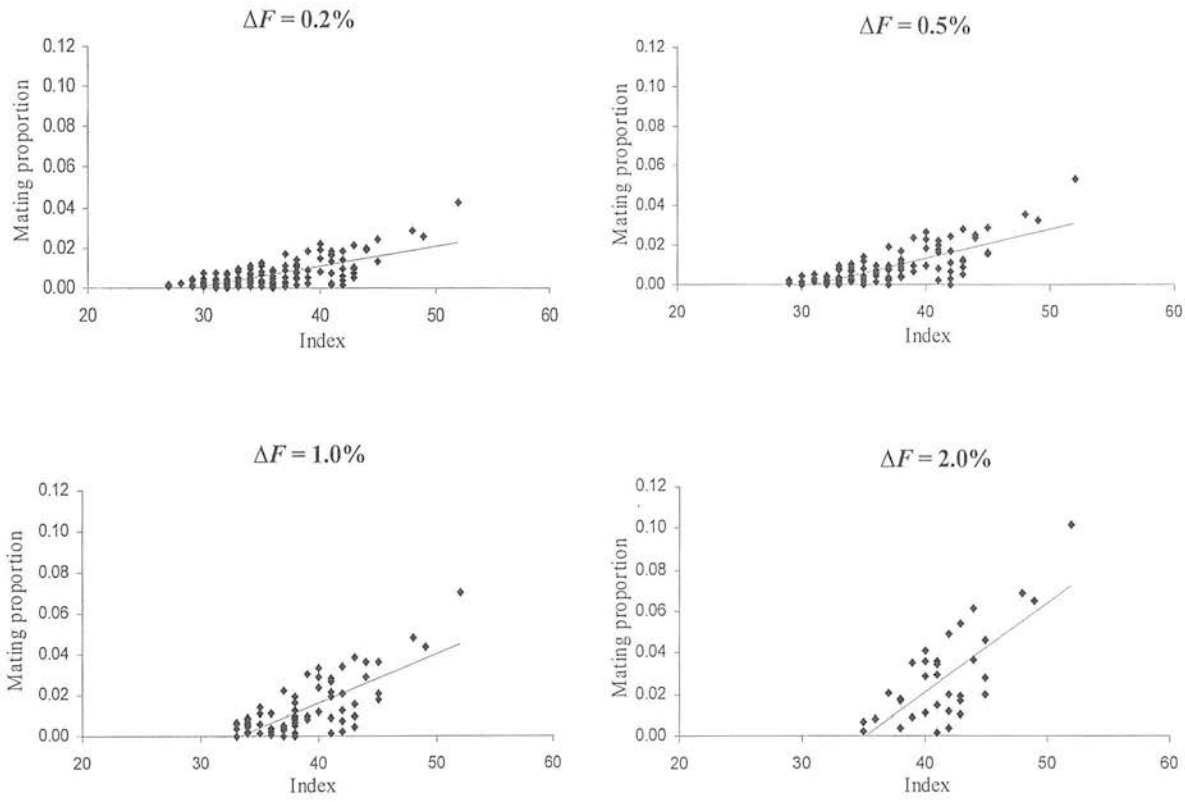


Figure 2.6. Relationship between optimised mating proportions and index scores of selected candidates for four levels of restriction on the rate of inbreeding (ΔF) when mating proportions of both sexes were optimised for Aberdeen Angus.



3. CHAPTER THREE

Mendelian Sampling Terms as the Selective Advantage in Optimum Breeding Schemes with Restrictions on the Rate of Inbreeding

3.1. Introduction

Dynamic selection algorithms for simultaneously managing genetic gain and inbreeding have been developed during the last decade (Wray and Goddard, 1994; Brisbane and Gibson, 1995; Meuwissen, 1997; Grundy *et al.,* 1998). The general framework can be described as a constrained quadratic optimisation problem on the usage of selection candidates. Although different objective functions have been used, they all have been of the form $\mathbf{c}^T\mathbf{g} - \lambda\mathbf{c}^T\mathbf{A}\mathbf{c}$, where $\mathbf{c}$ is the vector of mating proportions (usage) of selection candidates, $\mathbf{g}$ is the vector of estimated breeding values (*EBV*s) of selection candidates, $\mathbf{A}$ is the numerator relationship matrix among selection candidates and the factor λ represents a penalty on the increase in inbreeding. A key property of such objective function is that it keeps the method used for genetic evaluation and the restriction in inbreeding separated. Thus, *EBVs* can be estimated with the best available technique (i.e., Best Linear Unbiased Predictions or BLUP) and the selection policy is independently chosen according to risk preference. Wray and Goddard (1994) and Brisbane and Gibson (1995) implemented a procedure with fixed λ as a constraint on cumulative inbreeding. On the other hand, Meuwissen (1997) used a Lagrangian multiplier to constrain the average relatedness ($\mathbf{c}^T\mathbf{A}\mathbf{c}$) to be less or equal a pre-defined value as an attempt to restrict the rate of inbreeding (ΔF), rather than the absolute average inbreeding level. The algorithm of Meuwissen (1997) explicitly found near optimal solutions for maximising genetic gain (ΔG) under BLUP-EBV selection but the constraint used failed to keep a constant rate of inbreeding in medium and the long-term. Using different constraints, Grundy *et al.,* (1998a) achieved optimal solutions on a long-term context, allowing the inbreeding rate being constrained to specific values over generations. Both approaches give the same solution if the constraints of Grundy *et al.,* (1998a) are used in the objective function of Meuwissen (1997).

The simulation studies of Meuwissen (1997) and Grundy *et al.,* (1998a) showed that optimal selection can potentially achieve at least 20% higher genetic gains than traditional BLUP truncation selection at the same level of ΔF. Moreover, when

applied to real livestock populations of beef cattle and sheep, even higher gains can be expected if the optimisation outcome were to be followed (see Chapter 2).

The unified theory of genetic contributions put forward by Woolliams and Thompson (1994) has provided the necessary framework for linking long-term genetic contributions to rates of gain and inbreeding in livestock populations. The long-term genetic contribution (r_i) of an individual i was defined by Wray and Thompson (1990) as the proportion of genes it contributes in the long-term to the population. These authors demonstrated that the rate of inbreeding per generation is proportional to the sum of squares of long-term contributions, $E(\Delta F) = \frac{1}{4}\sum r_i^2$ with the sum taken over all ancestors. Woolliams and Thompson (1994) showed that the rate of genetic gain can be expressed in terms of the covariance between long-term contributions and the Mendelian sampling terms (a_i), $E(\Delta G) = \sum r_i a_i$. This definition of gain makes explicit that genetic gain arises from '*good*' ancestors contributing more genes and that the process of contribution of genes is of multi-generational nature. Furthermore, it makes explicit that sustained genetic gain relies on the exploitation of the Mendelian sampling variation, that is, the new genetic variation created each generation (Woolliams *et al.*, 1999).

There is no available framework for predicting ΔG under constrained ΔF, that is, in the context of quadratic indices. Although there have been notable developments on the deterministic predictions of ΔF for livestock populations undergoing mass selection (Wray and Thompson, 1990; Bijma *et al.*, 2000) and BLUP selection (Bijma and Woolliams, 2000; Bijma *et al.*, 2001) this work has been developed under truncation selection. Grundy *et al.*, (1998a) attempted the derivation of a deterministic prediction of ΔG under quadratic optimisation and constrained ΔF accounting for the lack of knowledge of r and a, but the approach was unclear and not formulated in a usable expression. In other words, the problem of predicting ΔG for a given constraint in ΔF and a set of known design variables (e.g., resources and trait heritability) has not yet been solved.

The study of Grundy *et al.*, (1998a) showed the link between the optimisation problem of maximising ΔG at fixed ΔF using quadratic indices and the management of long-term genetic contributions. These authors theoretically showed that the optimal solution could be achieved by linearly allocating long-term contributions of selection candidates according with their Mendelian sampling terms. A deterministic prediction of the potential ideal rate of gain (ΔG_{ideal}) that would be achievable with the ideal outcome of the constrained optimisation problem was derived in terms of the number of candidates per generation, heritability and desired ΔF.

The genetic responses obtained by Grundy *et al.*, (1998a) after applying the dynamic selection algorithm in stochastic simulations were always substantially lower than ΔG_{ideal} for a given heritability and inbreeding constraint. This was attributed to the lack of knowledge about r and a at selection time and to the fact that in a multi-generational process, long-term contributions of individuals cannot be independently managed, as pointed out by Woolliams *et al.*, (2002).

A required step towards the derivation of deterministic predictions of ΔG is the comprehensive understanding of the mechanics of quadratic indices, not yet addressed. The attempt of Grundy *et al.*, (1998a) of tackling the problem only considered the empirical relationship between r and *EBV*s, without explicit reference to its relevant relationship to a, clearly explicit in the expression for ΔG of Woolliams and Thompson (1994).

Under truncation selection, genetic gain can be predicted using simple linear regression theory and the predicted selection differential ($\hat{S}$) expressed as $\hat{S} = b_{g.I} i \sigma_I$, where $b_{g.I}$ is the regression of the true breeding value on the selection index used (e.g., BLUP), i is the selection intensity, and σ_I is the standard deviation of the selection index. On the other hand, under quadratic optimisation there is no explicit definition of the achievable $\hat{S}$ for a given constraint in ΔF.

The objective of this study was to demonstrate that quadratic optimisation manages individual contributions in relation to the best available information on the Mendelian sampling term rather than breeding values. This basic relationship has not yet been addressed and it is a necessary step required for the development of usable and tractable predictions of genetic gain for selection tools that maximizes progress with constrained rates of inbreeding. An empirical route was followed through the use of stochastic simulations and comparisons with traditional truncation selection.

3.2. Methods

3.2.1. Genetic model and simulation procedure

Selection over several generations was modelled using stochastic computer simulations. An additive infinitesimal model (Bulmer, 1971) was considered for the trait under selection. The true breeding values for animals on the base population were obtained from a normal distribution with mean zero and variance equal to the heritability (h^2) of the trait. Thus, the phenotypic variance (σ_p^2) was assumed equal to unity. In subsequent generations, the true breeding value of the progeny was obtained as half the sum of the true breeding values of their parents plus a random Mendelian sampling term. The Mendelian sampling term of an individual was sampled from a normal distribution with mean zero and variance $\sigma_a^2 = \frac{1}{2}\left[1 - \frac{1}{2}\left(F_S + F_D\right)\right]h^2$ where F_S and F_D are the inbreeding coefficients of the sire and dam, respectively. The phenotypic value for any individual in any generation was obtained by adding to the true breeding value, an environmental component sampled from a normal distribution with mean zero and variance 1- h^2. The genetic evaluation to estimate breeding values was carried out using a BLUP animal model. Populations with discrete generations were evaluated over 10 generations of selection. In the base generation (t=0), N (100 or 200) individuals ($N/2$ males and $N/2$ females) with family structure were generated. The first generation of selection was obtained from mating of animals selected at t=0. Mating amongst selected candidates was at random. A total of 100 replicates were performed.

3.2.2. Selection methods

3.2.2.1. Optimisation selection

The optimisation algorithm described by Meuwissen (1997) for obtaining maximum genetic gain while constraining the inbreeding rate to a specific value was used. This procedure relies on the maximisation of an objective function using Lagrangian multipliers to achieve the restriction on the rate of inbreeding and on the maximum contribution per sex (i.e., one half). The constant rate of inbreeding in the long-term was achieved by setting the constraint on the average coancestry of selected candidates ($\mathbf{c}^{\mathrm{T}}\mathbf{A}\mathbf{c}/2$) to $1-(1-\Delta F)^{t}$ where t is the generation number (Grundy *et al.*, 1998a). With optimal selection, the number of selected candidates and their mating proportions or contributions to the next generation are optimised. The output of this optimisation procedure is a vector of mating proportions ($\mathbf{c}$) of candidates at any particular generation. The optimal number of offspring for an individual i is $2Nc_i$ (a real number), and the actual (integer) number of offspring per parent was obtained following Grundy *et al.*, (1998a). Selected candidates are those with $c_i > 0$ and they will contribute to the next generation according to their c_i values.

3.2.2.2. Truncation selection

In standard truncation selection, a fixed number of candidates from each sex (those with the highest *EBVs*) were selected each generation. The population structure was chosen to enable comparisons with optimised selection at the same rate of inbreeding. An optimised scheme with a restriction of $\Delta F = 0.01$ per generation was first run, and then, a truncation scheme with a similar population structure was chosen. The desired ΔF under truncation selection was achieved after selecting 25 males and 25 females out of N=100 candidates per generation. The mating ratio of 1 was used since this was the typical optimal value found by Villanueva *et al.*, (1996, 1997) for small schemes. Mating pairs were randomly allocated and every female

had 2 offspring of each sex. Thus, in contrast with optimal selection, a fixed contribution to the next generation of 0.02 was allocated to each selected candidate.

3.2.3. Long-term genetic contributions and estimated Mendelian sampling term at selection time and at convergence

The relationships between long-term genetic contributions (r), estimated Mendelian sampling terms ($\hat{a}$) and *EBVs*, were studied for different values of h^2 (0.01, 0.25, 0.50 and 0.99) and three levels of constraint on the rate of inbreeding (0.005, 0.01 and 0.02). The long-term contribution of an ancestor i to a descendant j is defined as the proportion of genes of j that are expected to derive by descent from the ancestor i (Wray and Thompson, 1990; Woolliams *et al.*, 1999). The long-term genetic contributions of ancestors born in generation 3 to descendants born in generation 8 were calculated by tracing back the pedigree from descendants to ancestors using the algorithm described by Woolliams and Mäntysaary (1995). The term 'at convergence' used throughout the paper refers to generation 8. The estimated Mendelian sampling term for selection candidate i was obtained as $\hat{a}_i = EBV_i - [\frac{1}{2}(EBV_S + EBV_D)]$, where EBV_i, EBV_S, EBV_D are the BLUP estimated breeding values for the individual, its sire and its dam, respectively. The *EBV*s used to calculate $\hat{a}_i$ were either those at the time of selection (i.e., EBV_{sel} obtained at generation 3 to estimate $\hat{a}_{sel}$) or at the convergence of long-term contributions of selected candidates (i.e., EBV_{conv} obtained at generation 8 to estimate $\hat{a}_{conv}$).

3.2.4. Methodology for comparison between selection methods

The potential selective advantages based on $\hat{a}$ and EBV, and optimal and truncation selection were compared using linear regression methods. Multivariate regression of contributions, c and r, either on $\hat{a}_{sel}$ and EBV_{sel}, or on $\hat{a}_{conv}$ and EBV_{conv} for selected individuals (i.e., those with $c>0$) were performed. For instance, for the long-term contribution of an animal i the regression model was:

$$r_i = u + b_{r,\hat{a}_{conv}} \hat{a}_{conv_i} + b_{r,EBV_{conv}} EBV_{conv_i} + e_i$$

where u is the intercept, $b_{r,\hat{a}_{conv}}$ and $b_{r,EBV_{conv}}$ are partial regression coefficients on each selective advantage (i.e., $\hat{a}_{conv_i}$ and EBV_{conv_i}) and e_i is the residual term. For c, the analysis was performed with selective advantages estimated both at selection time and at convergence, whereas for r only selective advantages estimated at convergence were used. Hence, four partial regression coefficients involving c ($b_{c,\hat{a}_{sel}}$ and $b_{c,EBV_{sel}}$, and, $b_{c,\hat{a}_{conv}}$ and $b_{c,EBV_{conv}}$) and two involving r ($b_{r,\hat{a}_{conv}}$ and $b_{r,EBV_{conv}}$) on selective advantages were calculated. Simple regressions of contributions on selective advantages were also used for comparison with the multivariate approach (i.e., the same notation as above was used). The relationship between contributions and each of the two selective advantages was graphically illustrated by using simple linear relationships including correlation coefficients, and residual variances.

The residual error from the simple linear regression of r on $\hat{a}_{conv}$ has been postulated by Woolliams *et al.*, (2002) as a measure of departure from the theoretical ideal optimal solution proposed by Grundy *et al.*, (1998a), and relates to the proportion of $\frac{1}{4}\sum r^2$ (thus, ΔF) that does not generate gain efficiently. This concept is further developed in the discussion section. The parameter $\frac{1}{4}\sum r_{dev}^2$, where r_{dev} are the deviations of the observed r from the predicted values after regressing r on $\hat{a}_{conv}$ or on EBV_{conv} was calculated for a reference scheme of N=100, h^2=0.25 and ΔF constraints of 0.01 and 0.02 per generation and for both selection methods.

An additional comparison between quadratic optimisation and truncation selection was established from the distributions of the selective advantage of selected and unselected individuals. The frequency distributions of both groups were plotted after arbitrarily defining 32 bin classes for $\hat{a}_{sel}$ and 64 bin classes for EBV_{sel} and summary distribution statistics were calculated. For $\hat{a}_{sel}$ and EBV_{sel} under quadratic

optimisation, and for $\hat{a}_{sel}$ under truncation selection, there is an overlap in the frequency distribution of those selected and unselected candidates. The degree of overlap between distributions was measured as the probability that the selective advantage ($\hat{a}_{sel}$ or EBV_{sel}) of a randomly taken individual from the selected group was equal to or greater than the selective advantage of a randomly taken individual from the unselected group. For instance, for $\hat{a}_{sel}$ this probability was approximated as $p(\hat{a}_{i\_SELECT} \geq \hat{a}_{j\_UNSELECT}) = \sum_{k=1}^{\max bin}\left[p(\hat{a}_{i\_SELECT})_k \times \left(\sum_{l=1}^{bin_l = bin_k} p(\hat{a}_{j\_UNSELECT})_l \right) \right]$, where $p(\hat{a}_{i\_SELECT})_k$ is the probability of an individual i from bin k being selected, and $p(\hat{a}_{j\_UNSELECT})_l$ is the probability of an individual j from bin l being not selected. Likewise, the corresponding probability was calculated for EBV_{sel}.

3.3. Results

3.3.1. *Relationships of long-term genetic contributions and mating proportions with breeding values and Mendelian sampling terms*

The regression coefficient of the long-term genetic contributions on the mating proportions ($b_{r,c}$) of selected candidates born in generation 3, and the partial regression coefficient of these two on the estimated Mendelian sampling term ($b_{r,\hat{a}_{conv}}$ and $b_{c,\hat{a}_{conv}}$) and on the estimated breeding values ($b_{r,EBV_{conv}}$ and $b_{c,EBV_{conv}}$) at convergence are shown in Table 3.1. The $b_{r,c}$ was around one for all combinations of h^2 and ΔF, indicating that mating proportions (i.e., contributions to the next generation) are unbiased estimators of the corresponding long-term contributions as demonstrated by Grundy *et al.*, (1998a).

The most relevant result from Table 3.1 was that with quadratic optimisation the contributions of selected candidates are more related to the best estimates of the

Mendelian sampling term ($\hat{a}_{conv}$) than to the best estimate of the breeding value (EBV_{conv}). Although the same relationships were observed for both ΔF constraints, the partial regression coefficients were higher for $\Delta F = 0.02$ than $\Delta F = 0.01$. For all combinations of h^2 and ΔF, both $b_{r,\hat{a}_{conv}}$ and $b_{c,\hat{a}_{conv}}$ were much greater than the corresponding $b_{r,EBV_{conv}}$ and $b_{c,EBV_{conv}}$. Although in the majority of cases $b_{r,EBV_{conv}}$ and $b_{c,EBV_{conv}}$ were significant (i.e., see Table 3.1) their contribution to decreasing the residual mean of squares of the regression was always negligible. For all combinations of parameters in Table 3.1, the difference of the residual mean of squares from the multivariate model and that from a simple regression of r or c on $\hat{a}_{conv}$ was always in the order of 10^{-7} for $\Delta F = 0.01$ and 10^{-6} for $\Delta F = 0.02$. Hence, a simple regression of contributions on $\hat{a}_{conv}$ explained virtually the same proportion of the total variation in r or c as the multivariate regression model.

The partial regression coefficient $b_{r,\hat{a}_{conv}}$ was typically higher than $b_{c,\hat{a}_{conv}}$, although they became similar as h^2 increased (e.g., for ΔF= 0.01 and h^2= 0.99, $b_{r,\hat{a}_{conv}}$ was 0.027 and $b_{c,\hat{a}_{conv}}$ was 0.025). In contrast, $b_{r,EBV_{conv}}$ and $b_{c,EBV_{conv}}$, when significant, were always very similar. For a given ΔF constraint, all the regressions on selective advantage shown in Table 3.1 decreased in magnitude as h^2 increased from 0.01 to 0.99. This result was somewhat expected since as h^2 increased, the accuracy, hence the variance of the index (i.e., the scale of $\hat{a}_{conv}$) increased (the variance of r and the correlation between r and $\hat{a}_{conv}$ remained constant). For instance, for $b_{r,\hat{a}_{conv}}$ and $b_{c,\hat{a}_{conv}}$, and ΔF=0.01, the variance of the estimated Mendelian sampling term ($\sigma^2_{\hat{a}}$) among selected candidates increased from 4.6×10^{-5} for h^2=0.01 to 0.211 for h^2=0.99.

Under quadratic optimisation, the stronger relationship between contributions (either c or r) and the best estimate of the Mendelian sampling term holds at all times, that is, from selection time through to the convergence of r. This is clearly shown in Table 3.2, where it is observed that the partial regression coefficients of mating

proportions and long-term contributions on estimated Mendelian sampling term ($b_{c,\hat{a}_{sel}}$ and $b_{r,\hat{a}_{conv}}$, respectively) were much greater than the partial regression coefficients of mating proportions and long-term contributions on estimated breeding values ($b_{c,EBV_{sel}}$ and $b_{r,EBV_{conv}}$, respectively) for all values of h^2 and constraints on ΔF. As pointed out before, although the partial regression coefficients on EBV were significant, the contribution of this selective advantage, both at selection and at convergence, on reducing the residual mean of square of the regression was only marginal. The magnitude of the difference between the residual mean of squares from the multivariate model and that from a single linear regression of c on $\hat{a}_{sel}$ or r on $\hat{a}_{conv}$ always was within range from 10^{-5} to 10^{-8} for the whole range of h^2 and ΔF constraints.

The empirical relationship between long-term contributions and $\hat{a}_{conv}$ or EBV_{conv} is illustrated in Figure 3.1. The pooled correlation coefficients and the $\sum r_{dev}^2$ from the simple regression of r on $\hat{a}_{conv}$ and EBV_{conv} including all the observations generated across replicates are presented for h^2=0.25 and ΔF=0.01 or ΔF=0.02. Results indicate that the ultimate estimate of the Mendelian sampling term is a better predictor of the long-term contributions than the estimated breeding values. A much higher correlation coefficient between r and $\hat{a}_{conv}$ ($\rho_{r,\hat{a}_{conv}}$) compared to the corresponding correlation coefficient between r and EBV_{conv} ($\rho_{r,EBV_{conv}}$) was observed (Figure 3.1). Also, the correlation coefficients were greatest for the tightest ΔF restriction (ΔF=0.01).

The sum of squared deviations of the predicted r from the simple regression $b_{r,\hat{a}_{conv}}$ or $b_{r,EBV_{conv}}$ (i.e., along the straight lines in Figure 3.1) was used to calculate $\frac{1}{4}\sum r_{dev}^2$. This parameter was much lower when r was predicted from the ultimate estimate of the Mendelian sampling term than when predicted from the ultimate estimate of the breeding value (i.e., 0.000825 and 0.00295, respectively for ΔF =0.01). Also, for both selective advantages, $\frac{1}{4}\sum r_{dev}^2$ was lower for the tightest ΔF constraint. For $\hat{a}_{conv}$,

$\frac{1}{4}\sum r_{dev}^2$ was 0.000825 for ΔF =0.01 and 0.003260 for ΔF =0.02. These values represent 8.8% and 17.8% of $\frac{1}{4}\sum r^2$ (i.e., 0.009365 for ΔF =0.01 and 0.018242 for ΔF =0.02), respectively.

Having shown that $\hat{a}_{conv}$ is a better predictor of the long-term genetic contributions than EBV_{conv}, it is relevant to describe the relationship between the estimators of r, the mating proportions (c), and the estimated Mendelian sampling terms at the time of selection ($\hat{a}_{sel}$). Figure 3.2 shows these relationships for ΔF =0.01 or 0.02 and h^2 =0.25. The c was plotted against $\hat{a}_{sel}$ (i.e., at generation 3) and superimposed to the plot of r against $\hat{a}_{conv}$ (i.e., $b_{r,\hat{a}_{conv}}$ in Figure 3.1). For ΔF=0.01 the simple linear regressions $b_{c,\hat{a}_{sel}}$ and $b_{r,\hat{a}_{conv}}$ were similar (0.132 ± 6.72x10^{-4}and 0.129 ± 7.28x10^{-4}, respectively), whereas for ΔF=0.02, $b_{c,\hat{a}_{sel}}$ was significantly higher than $b_{r,\hat{a}_{conv}}$ (0.247 ± 2.74x10^{-3} and 0.213 ± 2.27x10^{-3}, respectively). A comparison between $b_{c,\hat{a}_{sel}}$ and $b_{r,\hat{a}_{conv}}$ for a broader range of ΔF constraints and h^2 is presented in Table 3.2. These results indicate that the outcome of the quadratic optimisation at the time of selection (i.e., $b_{c,\hat{a}_{sel}}$) is a reasonable estimator of the ultimate relationship between r and $\hat{a}_{sel}$ (i.e., $b_{r,\hat{a}_{conv}}$). This was the case for the whole range of ΔF constraints and in particular when the trait h^2 is not too low (e.g., 0.01).

3.3.2. *Estimated Mendelian sampling terms of selected and unselected candidates under Quadratic Optimisation*

The frequency distribution of the estimated Mendelian sampling at the time of selection ($\hat{a}_{sel}$) of selected and unselected candidates born in generation 3 across the 100 replicates for N=100 (i.e., 10,000 observations) for both levels of ΔF constraint is presented in Figure 3.3. For ΔF=0.01, the proportion of selected candidates (p) was 0.5943. For ΔF=0.02, the corresponding p was reduced to 0.3584. The mean $\hat{a}_{sel}$ was

higher for selected candidates than for unselected candidates, and the difference in $\hat{a}_{sel}$ between both groups was constant for both ΔF constraints (i.e., about 0.56 standard deviations of the true Mendelian sampling term in the base population). For the most relaxed constraint (ΔF=0.02) the within group variance of $\hat{a}_{sel}$ among unselected candidates was greater than among selected candidates (0.0081 and 0.0055, respectively).

The coefficient of right skewness of the distribution of $\hat{a}_{sel}$ for selected candidates was about 0.7 for both levels of inbreeding constraint. On the other hand, the degree of left skewness of the distribution of non-selected candidates changed with the ΔF restriction (–0.77 for ΔF=0.01 and -0.58 for ΔF=0.02). Thus, a greater ΔF resulted in a more centred distribution of those individuals that remained unselected.

3.3.3. Qualitative comparison between optimal selection and truncation selection

A multivariate regression analysis of r on the ultimate estimates of the Mendelian sampling terms and breeding values was performed for h^2=0.25 and ΔF=0.01. In contrast to quadratic optimisation, under truncation selection both estimated partial regressions on selective advantage were significant (p<0.01) and of the same magnitude (i.e., $b_{r,\hat{a}_{conv}} = 0.06$ and $b_{r,EBV_{conv}} = 0.04$). In fact, it was EBV_{conv} the term that appeared more relevant in explaining the variance of the long-term contributions. A single regression model of r on EBV_{conv} gave virtually the same mean squared error than the multivariate regression approach (i.e., 3.4×10^{-4} and 3.8×10^{-4}, respectively).

The relationship between long-term contributions and estimated Mendelian sampling terms or estimated breeding values at convergence of r (i.e., generation 8) across replicates are shown for truncation selection in Figure 3.4 for h^2=0.25 and ΔF=0.01. Under truncation selection, both selective advantages were equally good predictors of the long-term contributions. Both the correlation between r and $\hat{a}_{conv}$, and the

correlation between r and EBV_{conv} were around 0.6 (Figure 3.4), whereas under quadratic optimisation $\rho_{r,\hat{a}_{conv}}$ was 0.92 and $\rho_{r,EBV_{conv}}$ was 0.65 (Figure 3.1).

A key qualitative difference between both selection methods arose from the proportion of the sum of squares of long-term contributions that is converted to genetic gain. In contrast to quadratic optimisation, under truncation selection a much higher proportion of ΔF does not generate gain efficiently. The parameter $\frac{1}{4}\sum r_{dev}^2$ from the simple regression $b_{r,\hat{a}_{conv}}$ was 0.005 thus representing a proportion of 0.50 of $\frac{1}{4}\sum r^2$ for ΔF restricted to 0.01 per generation.

The comparison between the frequency distribution of $\hat{a}_{sel}$ and EBV_{sel} for selected and non-selected candidates under quadratic optimisation and truncation selection across generations and replicates is presented in Figure 3.5 for ΔF=0.01. The frequency distributions of $\hat{a}_{sel}$ with quadratic optimisation was already presented in Figure 3.3 but it was also included here to help the comparisons between selection methods. It should be noticed that the overlap between the frequency distributions of EBV_{sel} for selected and unselected candidates for truncation selection is purely an artefact created from the use of pooled data across replicates, since in a single replicate every selected candidate will have higher EBV_{sel} than every unselected candidate. However, this is not the case for the frequency distributions of $\hat{a}_{sel}$ where even under truncation selection, some overlapping can occur between selected and unselected candidates. For quadratic optimisation the overlapping between selected and unselected candidates is a property of the method and can occur for both $\hat{a}_{sel}$ and the EBV_{sel} on a single replicate or across replicates.

The degree of overlap between frequency distributions of selected and unselected candidates was different for quadratic optimisation and truncation selection and depended on the selective advantage used by the selection method. For quadratic optimisation $p(\hat{a}_{i\_SEL} \geq \hat{a}_{j\_UNSEL})$ was 0.988 while $p(EBV_{i\_SEL} \geq EBV_{j\_UNSEL})$ was 0.839. In contrast, for truncation selection, (i.e., using pooled observations)

$p(\hat{a}_{i\_SEL} \geq \hat{a}_{j\_UNSEL})$ was 0.832, whereas $p(EBV_{i\_SEL} \geq EBV_{j\_UNSEL})$ was 0.955. Thus, at selection time, the probability that the selective advantage of a selected candidate will be equal or greater than that of an unselected candidate was higher when the selective advantage was $\hat{a}_{sel}$ under quadratic optimisation and when it was EBV_{sel} under truncation selection. Other definitions of this probability (e.g., excluding the same selective advantage bin, thus, $p(\hat{a}_{i\_SEL} \rangle \hat{a}_{j\_UNSEL})$ made no qualitative difference to the result presented.

The difference between the mean $\hat{a}_{sel}$ of selected and non-selected individuals was greater under quadratic optimisation than under truncation selection (i.e., $0.56\sigma_a$, and $0.43\sigma_a$ respectively). In contrast, the difference between the EBV_{sel} mean of selected and non-selected individuals was greater under truncation selection than under quadratic optimisation (i.e., $1.11\sigma_a$ and $0.78\sigma_a$, respectively). Interestingly, the selection differential (S) on EBV_{sel} was higher under truncation selection than under quadratic optimisation (i.e., $0.19\sigma_p$ and $0.12\sigma_p$, respectively), whereas no substantial difference in S on $\hat{a}_{sel}$ (i.e., around $0.08\sigma_p$) between the groups of selected and unselected candidates was observed between both methods. The number of selected candidates was slightly higher for quadratic optimisation (27 males and 27 females) than for truncation selection (25 males and 25 females). Therefore, the higher ΔG observed for quadratic optimisation than that of truncation selection (0.234 and 0.206, respectively at generation 4) must arise from the allocation of greater mating proportions to individuals of higher selective advantage, as showed in Figure 3.2. On the other hand, under truncation selection all selected candidates are allocated the same mating proportion.

3.4. Discussion

This study has empirically demonstrated that in the context of quadratic optimisation, the selection advantage of selected candidates is a function of the estimated

Mendelian sampling term. We have shown that from initial selection to the convergence of long-term genetic contributions, quadratic optimisation attempts to allocate contributions of selected candidates according to the best available information on their estimated Mendelian sampling term, not their breeding value. Therefore, under quadratic optimisation, the selective advantage of selection candidates can be better described in terms of $\hat{a}$ rather than *EBV*, hence, candidates are selected or culled according with the estimate of their independent and unique superiority or inferiority with respect to the parental average. This is in contrast with BLUP truncation selection, where the fate of candidates depends upon their *EBVs*, hence their selective advantage is not independent from the parental average.

Grundy *et al.,* (1998a) postulated the relationship between long-term contributions and Mendelian sampling term under constrained inbreeding from a theoretical standpoint, and stated that the optimal solution for the optimisation problem could be obtained by an exact linear allocation of long-term contributions of selected candidates according to their Mendelian sampling term. This theoretical optimal solution is presented in Figure 3.6 (after Woolliams *et al.,* 2002) for a tight and a relaxed inbreeding constraint. Under the ideal outcome, a candidate will have a long-term contribution greater than zero only if its a is greater than the value (u) at the intercept of the regression line of r on a, and its converged contribution will be given by the regression equation (i.e., for a candidate i, $r_i = b(a_i-u)$, where b is the regression slope). For a tight ΔF constraint both u and b will be lower than for a more relaxed constraint.

Although Grundy *et al.,* (1998a) hypothesed the theoretical outcome of the optimisation, the present study is the first presenting empirical evidence of such a relationship. Grundy *et al.,* (1998a) concentrated their empirical analysis on demonstrating that quadratic optimisation was able to constraint ΔF to the desired value in the long-term, and on the relationship between mating proportions at selection time (c) and *EBVs*. On the other hand, in their simulation study, they did not make an explicit reference to the relationship between r and *EBV*, or more importantly, to the underlying relationship between r and $\hat{a}$.

The primary evidence that under quadratic optimisation the selective advantage is the Mendelian sampling comes from the result that the relationship between contributions (i.e., either c or r) and the estimated Mendelian sampling term ($\hat{a}_{conv}$) is greater than between contributions and estimated breeding value (EBV_{conv}). Not only were the partial regression coefficients involving $\hat{a}_{conv}$ much greater than those involving the EBV_{conv} irrespective of the ΔF constraint, but also, a simple regression of contributions on $\hat{a}_{conv}$ will perform as well as the multivariate regression on both selective advantages. Also, the correlation of r on selective advantages indicated that the estimated Mendelian sampling term was a better predictor of the long-term contributions. For both ΔF constraints, the correlation between r and $\hat{a}_{conv}$ ranged from 0.8 to 0.9 whereas the correlation between r and EBV_{conv} was around 0.6. In contrast, under truncation selection it was evident that there is not such a clear distinction between selective advantages on the management of contributions. Both partial regressions of r on $\hat{a}_{conv}$ or EBV were about 0.6, and also, there was no difference between both selective advantages as predictors of long-term contributions as the corresponding correlation coefficients were 0.6. Further evidence of the qualitative difference between quadratic optimisation and truncation selection arose from the probability that a selected candidate had an equal to or greater selective advantage than an unselected candidate. Under quadratic optimisation, the probability that at selection time a selected candidate will have an equal to or greater $\hat{a}_{sel}$ than that of an unselected candidate was higher than the corresponding probability under truncation selection. On the other hand, the opposite was true for EBV_{sel}, thus indicating that both methods work on different underlying selective advantages.

The most obvious qualitative difference between truncation selection and quadratic optimisation arises from the usage of selected candidates. While in truncation selection all individuals with EBV above a certain truncation point are allocated the same mating proportions, in quadratic optimisation the usage is allowed to vary

among selected candidates. Toro and Nieto (1984) first proposed the idea of allowing unequal contributions of selected candidates as an alternative to truncation selection under mass selection. By using a quadratic optimisation approach they minimised the sum of squares of contributions of selected candidates while maintaining a pre-defined selection intensity. This approach allowed the maximisation of the effective population size (judged by the mating proportions), by allocating optimal contributions according to ranking scores. The allocation of higher contributions to individuals with higher selective advantage is indeed a property of quadratic optimisation with pre-defined levels of ΔF.

The expression for genetic gain of Woolliams and Thompson (1994) made explicit that sustained genetic gain arises from creating a covariance between the r and a, (i.e., $E(\Delta G) = \sum r_i a_i$) hence exploiting the Mendelian sampling term variance (i.e., the new genetic variation created each generation). It is reasonable therefore, that an algorithm for maximising gain should relate contributions to the Mendelian sampling terms or at least to the their best available estimate. Figures 3.1 and 3.2 explicitly show that quadratic optimisation attempts to maximise this covariance from selection time to the convergence of the long-term contributions.

Since the true Mendelian sampling term of selected candidates is unknown, the quadratic optimisation relies on the use of their best estimates at the time of selection. Thus, at selection time contributions to the next generation are allocated according to their initial a estimates as pointed out by Woolliams *et al.,* (2002). Therefore, the accuracy of the estimate of the Mendelian sampling term (ρ_a) is one of the relevant parameters in determining the degree of 'noise' in the optimisation system, that is, the amount of departure from the ideal solution represented in Figure 3.6. The departures from the ideal solution, represented by the scatter around the straight lines in Figure 3.1, have been regarded by Woolliams *et al.,* (2002) as unavoidable 'contribution errors'. A quarter of the sum of squares of these deviations ($\frac{1}{4}\sum r_{dev}^2$) is an indication the proportion of $\frac{1}{4}\sum r^2$ that does not generate gain efficiently, and of course are a quarter of the residual sum of squares after regressing

r on $\hat{a}_{conv}$. Given that $E(\Delta F) = \frac{1}{4}\sum r_i^2$, the efficiency of the breeding scheme can be summarised by $1 - \frac{\frac{1}{4}\sum r_{dev}^2}{\Delta F}$. Figure 3.1 provided benchmark efficiency values of 0.917 and 0.837 for schemes constrained to ΔF of 0.01 and 0.02, respectively for h^2=0.25. These are indeed high efficiencies, although suggest that schemes with more relaxed constraints (i.e., higher ΔF) may be less efficient in the exploitation of genetic variation even when using quadratic indices. This indicates that the degree to which ΔF is constrained controls the amount of departure from the ideal solution, which was expected as the theoretical ideal outcome was derived under a constrained assumption. In contrast, under truncation selection an efficiency of only 0.50 for ΔF of 0.01 and h^2=0.25 was found.

We have shown that for all ΔF constraints and for a broad range of heritabilities (Figure 3.2 and Table 3.2), the outcome of the quadratic optimisation at selection time, summarised by $b_{c,\hat{a}_{sel}}$, is a good estimator of the optimisation outcome that would be obtained if converged long-term contributions and the best estimate of the Mendelian sampling term ($b_{r,\hat{a}_{conv}}$) were available. Greater deviations from the ideal outcome observed for the relationship of r on $\hat{a}_{conv}$ reduces the ability to manage individual contributions independently in a dynamic multiple generational selection process. As stressed by Woolliams *et al.* (2002), the change of an individual's contribution will affect the contribution of all its ancestors, and changing the contribution of a male or female parent through its offspring will affect the contribution of its mates. Therefore, although the estimate of the Mendelian sampling term will be more accurate as information accumulates, the management of an individual contribution becomes less and less independent as selection progresses.

An alternative way of assessing the impact of ρ_a on the outcome of the optimisation process is by observing the change in the 'effective number of parents' from selection to convergence of the long-term contributions. The sum of squares of mating proportions and the sum of squares of the long-term contributions of selected candidates can be regarded as 'effective number of parents' of equal contribution

(e.g., Roberston, 1965) at selection time, $N_c = \left[\sum c^2\right]^{-1}$, and at convergence, $N_r = \left[\sum r^2\right]^{-1}$, respectively. For instance, ΔF=0.01 and h^2=0.01, N_c was 43.2 whereas N_r was 26.7, and ρ_a increased from 0.052 at selection time to 0.103 at convergence. On the other hand, for h^2=0.99, N_c was 33.1 whereas N_r was 26.4, and ρ_a was 0.968 both at selection time and at convergence. This indicates that for a given ΔF constraint the process of building up of ρ_a (thus, the agreement between the initial and the converged solution) is controlled by the trait h^2.

Our results provide the relevant empirical validation for the theoretical optimal solution of Grundy *et al.* (1998a) for the optimisation problem under constrained ΔF formulated in terms of r and a. It should be emphasized that our observations on the outcome and mechanics of quadratic optimisation are drawn from an ongoing use of dynamic selection in the breeding scheme, and the same outcome might not be achieved after only a single generation application of the method. Importantly, this study clearly provides evidence that the outcome of the quadratic optimisation relies on the accuracy of the knowledge of the Mendelian sampling terms at the time of selection. Therefore, there is a need for developing such deterministic predictions for the accuracy of predicted Mendelian sampling terms in order to obtain predictions of gain in a constrained optimisation context.

The use of Mendelian indices was proposed by Woolliams and Thompson (1994) as a way of explicitly altering the amount of weight given to family information for the flexible management of genetic gain and inbreeding. The index proposed relied on the decomposition of estimated breeding values and the re-weighing of the estimated Mendelian terms, progressively reducing weights given to more distant ancestors. This was examined by Grundy *et al.,* (1998b), and found to be useful but difficult to generalise. Whilst the concept of using Mendelian sampling terms as a selective advantage was in the appropriate direction, the implementation of the concept is most effectively done through the quadratic index rather than the linear index described by Grundy *et al.* (1998b).

3.5. Tables

Table 3.1. Regression of long-term contributions on mating proportions ($b_{r,c}$), partial regression coefficients from a multivariate analysis of long-term contributions and mating proportions on estimated Mendelian sampling terms and estimated breeding values ($b_{r,\hat{a}_{conv}}$ and $b_{r,EBV_{conv}}$ and, $b_{c,\hat{a}_{conv}}$ and $b_{c,EBV_{conv}}$, respectively) at convergence of the long-term contributions, for a range of heritabilities (h^2) and two constraints on the rate of inbreeding (ΔF). Mating proportions were taken at generation 3 while the long-term contributions were calculated for candidates (ancestors) born at generation 3 to descendants born in generation 8 (i.e., convergence). Only selected candidates were included and N was 100 candidates per generation [a, b].

ΔF	h^2	$b_{r,c}$	$b_{r,\hat{a}_{conv}}$	$b_{r,EBV_{conv}}$	$b_{c,\hat{a}_{conv}}$	$b_{c,EBV_{conv}}$
0.01	0.01	1.003	2.405	0.036	0.689	0.037
	0.25	1.015	0.126	0.002	0.053	0.002
	0.50	1.026	0.071	0.000	0.040	0.001
	0.99	1.054	0.027	0.001	0.025	0.001
0.02	0.01	0.962	3.620	0.004	0.699	0.027
	0.25	0.970	0.201	0.009	0.060	0.008
	0.50	1.002	0.114	0.006	0.052	0.006
	0.99	1.018	0.058	0.002	0.056	0.002

[a] Pooled regression analysis were performed over 10 generations of selection and 100 replicates for selected candidates.

[b] All regression coefficients were significant ($p<0.01$) apart from $b_{r,EBV_{conv}}$ and $b_{c,EBV_{conv}}$ which were non-significantly different from zero ($p>0.05$) for ΔF=0.01 and h^2=0.50, for ΔF=0.02 and h^2=0.01 and for ΔF=0.02 and h^2=0.99.

Table 3.2. Partial regression coefficients from a multivariate regression of mating proportions on estimated Mendelian sampling term ($b_{c,\hat{a}_{sel}}$) and on estimated breeding value ($b_{c,EBV_{sel}}$) at selection time, and of long-term genetic contributions on estimated Mendelian sampling term ($b_{r,\hat{a}_{conv}}$) and on estimated breeding value ($b_{r,EBV_{conv}}$) at convergence of the long-term genetic contributions for a range of heritabilities (h^2) and two constraints on the rate of inbreeding (ΔF). Mating proportions were taken at generation 3 while the long-term contributions were calculated for candidates (ancestors) born at generation 3 to descendants born in generation 8 (i.e., convergence). Only selected candidates were included and N was 200 candidates per generation [a, b].

		Selection		*Convergence*	
ΔF	h^2	$b_{c,\hat{a}_{sel}}$	$b_{c,EBV_{sel}}$	$b_{r,\hat{a}_{conv}}$	$b_{r,EBV_{conv}}$
0.005	0.01	1.368	0.056	1.195	0.034
	0.25	0.064	0.001	0.063	0.002
	0.50	0.034	0.001	0.036	0.001
	0.99	0.013	0.001	0.013	0.001
0.01	0.01	2.296	0.116	1.848	0.051
	0.25	0.117	0.007	0.102	0.007
	0.50	0.066	0.003	0.061	0.003
	0.99	0.029	0.001	0.030	0.002
0.02	0.01	3.342	0.107	2.746	0.023
	0.25	0.190	0.017	0.153	0.013
	0.50	0.110	0.011	0.091	0.010
	0.99	0.065	0.002	0.070	0.003

[a] Pooled regression analysis were performed over 10 generations of selection and 100 replicates for selected candidates.

[b] All regression coefficients were significant ($p<0.01$).

3.6. Figures

Figure 3.1. Relationship between long-term contributions (r) and estimated Mendelian sampling term ($\hat{a}_{conv}$) or estimated breeding values (EBV_{conv}) at the convergence of the long-term contributions under quadratic optimisation for h^2=0.25 and two levels of ΔF constraint per generation. The correlation coefficients ($\rho_{r,\hat{a}_{conv}}$ and $\rho_{r,EBV_{conv}}$) and the sum of squares of residuals of long-term contributions ($\sum r_{dev}^2$) are presented. The population size is 100 candidates per generation and data is pooled across 100 replicates.

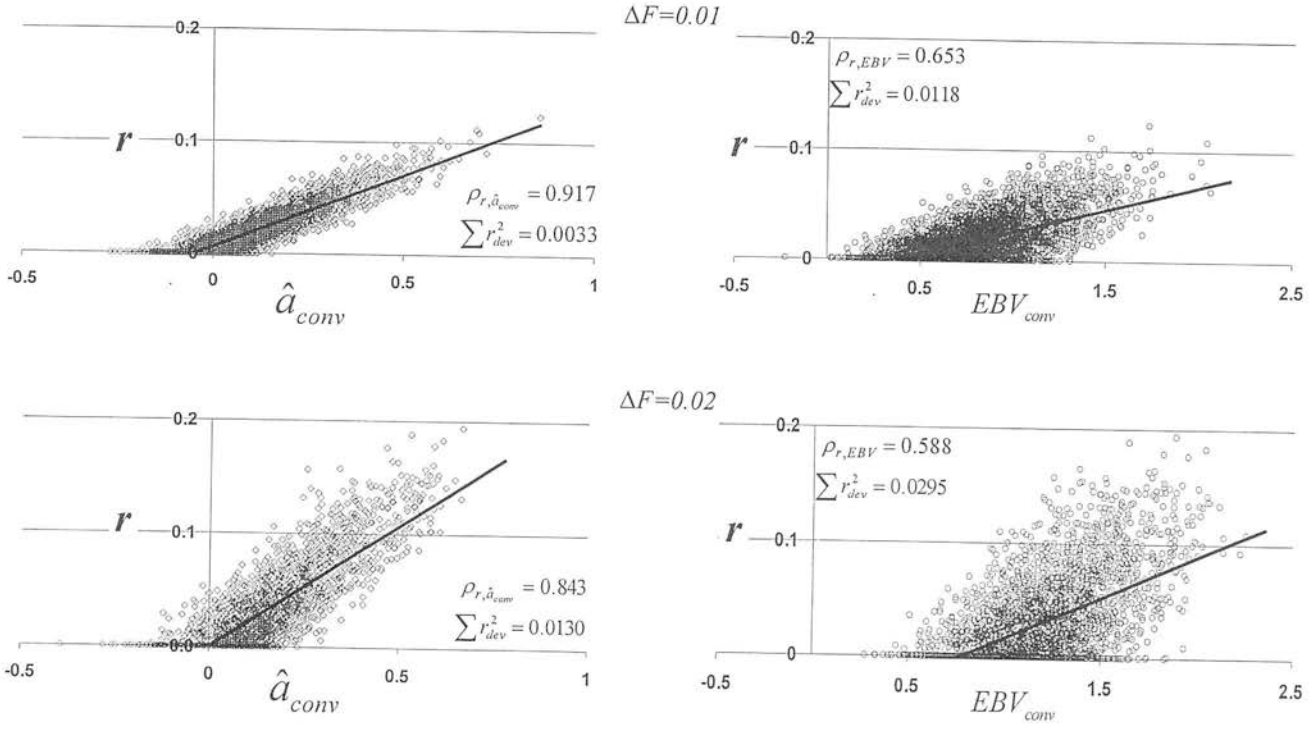


Figure 3.2. Relationship between long-term contributions (r) and the estimated Mendelian sampling term at convergence ($\hat{a}_{conv}$) (solid line) and between mating proportions (c) and estimated Mendelian sampling term at selection time ($\hat{a}_{sel}$) (broken line) under quadratic optimisation for h^2=0.25 and two levels of ΔF constraint per generation. The population size is 100 candidates per generation and data is pooled across 100 replicates.

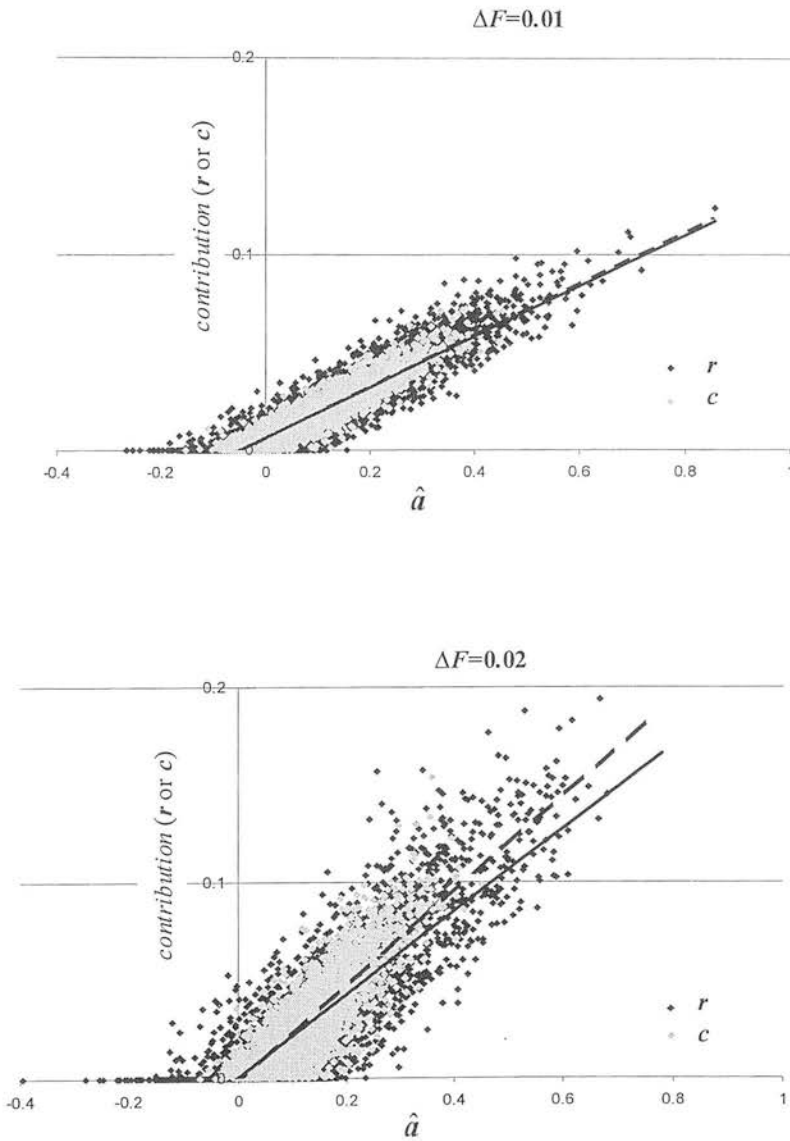
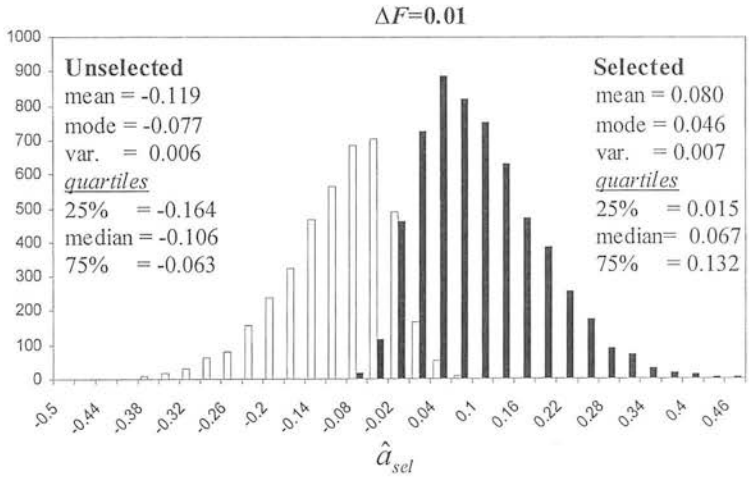


Figure 3.3. Frequency distribution of the Mendelian sampling at selection time ($\hat{a}_{sel}$) for selected and unselected candidates under quadratic optimisation for h^2=0.25 and two levels of ΔF per generation. The population size is 100 candidates per generation and data is pooled across 100 replicates.



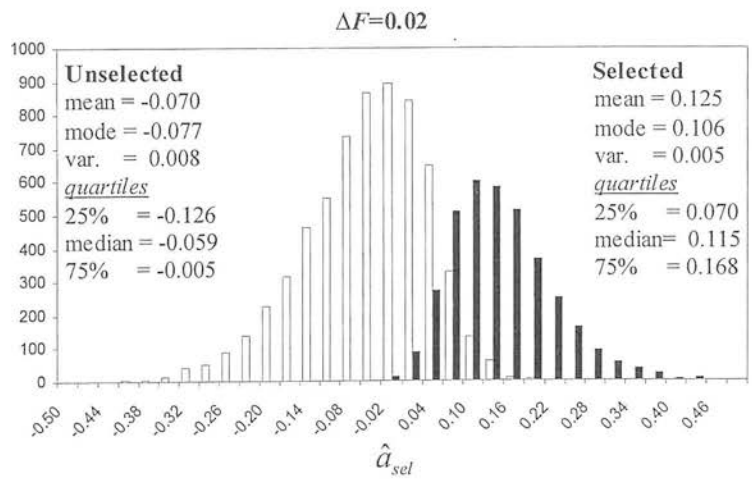
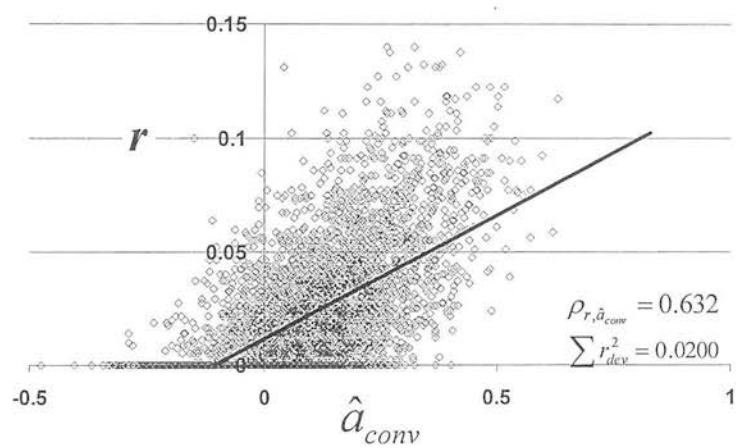


Figure 3.4. Relationship between long-term contributions (r) and estimated Mendelian sampling term ($\hat{a}_{conv}$) and estimated breeding values (EBV_{conv}) at convergence of the long-term contributions under truncation selection for h^2=0.25 and ΔF constrained to 0.01 per generation. The correlation coefficients ($\rho_{r,\hat{a}_{conv}}$ and $\rho_{r,EBV_{conv}}$) and the sum of squares of residuals of long-term contributions ($\sum r_{dev}^2$) are presented. The population size is 100 candidates per generation and data is pooled across 100 replicates.



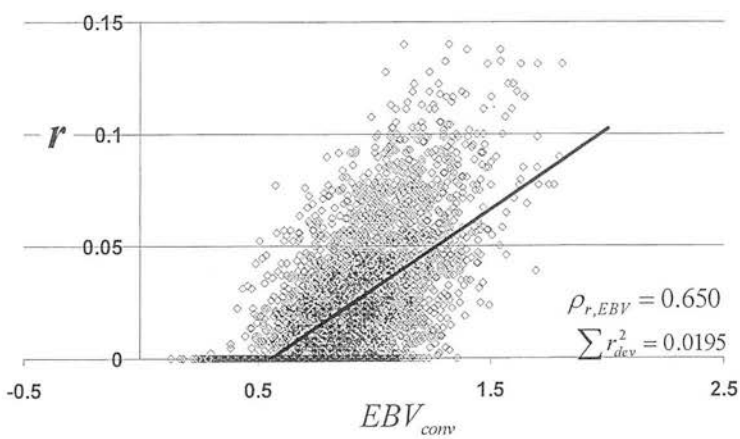


Figure 3.5. Frequency distribution of the estimated Mendelian ($\hat{a}_{sel}$) and the estimated breeding value at selection time (EBV_{sel}) for selected and unselected candidates under quadratic optimisation and truncation selection for h^2=0.25 and two ΔF constraints per generation. The population size is 100 candidates per generation and data is pooled across 100 replicates.

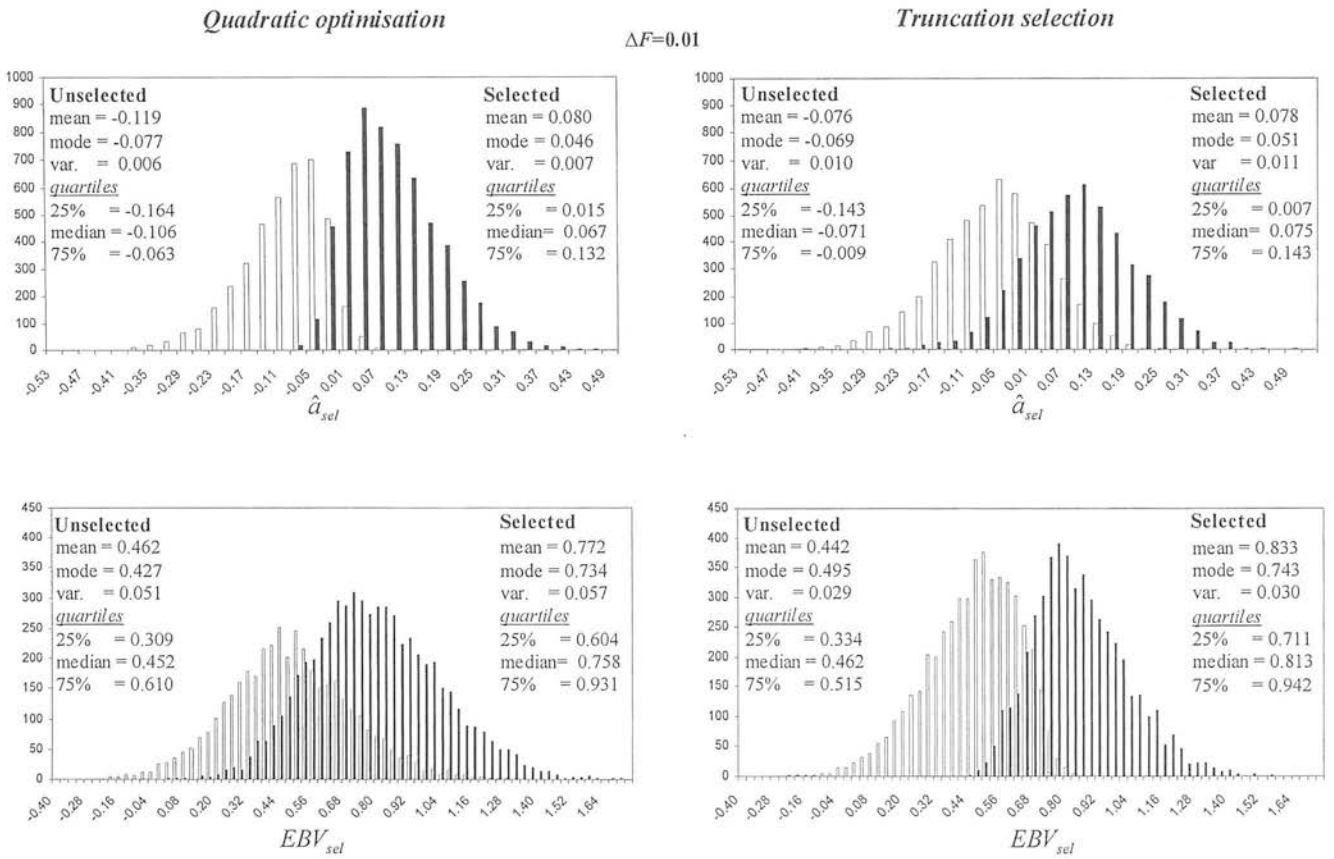
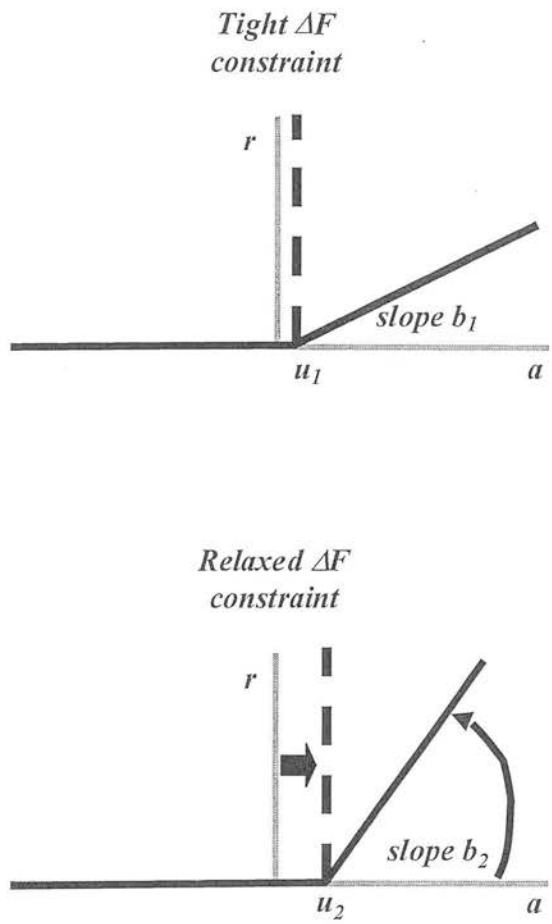


Figure 3.6. Ideal outcome for the optimisation of genetic gain for constrained rate of inbreeding (ΔF) according with Grundy *et al.*, (1998a), adapted from Woolliams *et al.*, (2002).



4. CHAPTER FOUR

Prediction of Rates of Genetic Gain from Quadratic Indices with Constrained Rates of Inbreeding

4.1. Introduction

Two different but complementary approaches have been followed to solve the problem of optimising breeding schemes for achieving the maximum rate of genetic gain (ΔG) at a pre-defined rate of inbreeding (ΔF). The first approach, fully deterministic, was aimed at solving the problem at the design stage of the breeding scheme and developed for mass (Villanueva *et al.*, 1996) and index selection (Villanueva and Woolliams, 1997) with discrete generations and for schemes for mass selection with overlapping generations (Villanueva *et al.*, 2000). This requires deterministic predictions of ΔG and ΔF for a set of design variables, such as the number of candidates per generation and trait heritability (h^2). The approach followed was to maximise a single objective function $\Phi = \Delta G - \lambda \Delta F$ that combined both deterministic predictions of ΔG and ΔF. The maximisation of this target function provided the optimum number of sires and dams selected given a fixed number of candidates and trait h^2. In the case of index selection not only the numbers selected, but also the weights given to the family information were optimised. These deterministic studies provided practical insight on the optimum design of breeding schemes with constrained ΔF under truncation selection (i.e., candidates are ranked and selected according to certain estimated breeding value threshold and contribute equally to the next generation).

The second approach for optimising breeding schemes with constrained ΔF refers to the routine operation of the breeding scheme and addresses the problem of which candidates should be selected and how much they should contribute to the following generation for maximising gain. This problem has been solved for scenarios with discrete and overlapping generations by using quadratic indices (Meuwissen, 1997; Grundy *et al.*, 1998a; Meuwissen and Sonesson, 1998; Grundy *et al.*, 2000). Under stochastic simulations, a broad range of extra response going from 20% to 60% from quadratic indices over truncation selection at the same ΔF has been observed (e.g., Meuwissen, 1997). Quadratic indices are beginning to be applied in real livestock populations (e.g., Weigel and Lin, 2002, in dairy cattle, and Chapter 2 in this thesis, in beef cattle and sheep). In this thesis (i.e., Chapter 2) substantial expected increases

in ΔG of 30% for beef cattle and 17% for sheep from quadratic indices when compared to traditional truncation selection at the same ΔF have been found.

Although the benefits in terms of gain from quadratic optimisation and BLUP selection are clear from empirical studies (Meuwissen, 1997; Chapter 2 and 3 in this thesis), there is not yet a framework available for obtaining deterministic predictions of ΔG while constraining ΔF with this type of approach. Therefore, while quadratic indices provide explicit optima for constrained ΔF, it is not possible to predict the benefits with respect to truncation selection nor how these are influenced by parameters such as the scheme size and trait h^2.

Grundy *et al.*, (1998a) theoretically showed that the ideal optimal solution for a given constraint in ΔF can be obtained after an exact linear allocation of long-term genetic contributions of selected candidates to their Mendelian sampling terms. Hence, providing the link between the optimisation of breeding schemes using quadratic indices and the maximisation of the covariance between long-term contributions (r) and Mendelian sampling terms (a) implicit in the definition of genetic gain of Woolliams and Thompson (1994) (i.e., $E[\Delta G] = \sum r_i a_i$). Grundy *et al.* (1998a) provided a deterministic prediction for the ideal rate of genetic gain (ΔG_{ideal}) assuming an exact allocation of long-term genetic contributions to Mendelian sampling terms. However, they found that the empirical gains obtained after applying quadratic indices were consistently lower than ΔG_{ideal}. The inability of quadratic optimisation to attain the ideal optimal solution arise from two sources: the lack of knowledge about the true Mendelian sampling term at selection time, and the inability of individually managing long-term contributions in a multi-generational process (Woolliams *et al.*, 2002).

The close relationship between long-term contributions and Mendelian sampling terms in breeding schemes using quadratic optimisation was empirically confirmed in the simulation study in Chapter 3 of this thesis. By using stochastic simulations, it was found that from selection to convergence of long-term contributions, quadratic

optimisation allocates contributions of selected candidates according to the best information on their Mendelian sampling term and not on their breeding values. Thus, under quadratic optimisation the selective advantage is a function of the estimated Mendelian sampling term, that is, the estimate of their independent and unique superiority or inferiority with respect to the parental average. In Chapter 3 it was also shown that although quadratic optimisation at all times linearly allocates higher long-term contributions to candidates with higher estimated Mendelian sampling terms there are always departures from the ideal solution. This study identified the role of the accuracy of the ultimate Mendelian sampling term estimate at the convergence of long-term contributions as an important factor for determining the degree of departure from the ideal optimal constrained solution of Grundy *et al.* (1988a).

The aim of this study was to develop a deterministic framework for predicting the rate of genetic gain under quadratic optimisation when selection is based upon BLUP estimates of breeding values. The approach was based on the development of deterministic predictions of the accuracy of the estimated Mendelian sampling term when long-term contributions converge and on the prediction of rate of ideal genetic gain previously developed by Grundy *et al.* (1998a). This will fill the gap in prediction tools for the design of breeding schemes under constrained rates of inbreeding and provides the necessary accompanying tool to the current available operational dynamic selection algorithms.

4.2. Methods

4.2.1. Notation

The description of the main notational conventions used throughout the chapter is presented in Figure 4.1.

4.2.2. *Ideal and upper bound genetic gain under constrained rate of inbreeding*

Grundy *et al.* (1998a) showed that the theoretical ideal solution for maximising the rate of gain (ΔG) when the rate of inbreeding (ΔF) is constrained to pre-defined values, could be obtained by exactly allocating the long-term genetic contribution (r) of selected candidates according to their Mendelian sampling terms (a). Although the derivation of the expression for the ideal rate of gain of Grundy *et al.* (1998a) was consistent in its own right it was somewhat ambiguous, hence a re-ordering of their approach is provided here. Grundy *et al.* (1998a) drew a parallel with the problem of Bondesson (1989) for the optimal allocation of clonal tree proportions (k_i) of known genetic values (A_i) to maximise gain with a constraint in genetic diversity (γ). They formulated the problem as:

$$\text{Maximise } \sum k_i A_i \text{ with the constraint } \sum k_i^2 \leq \gamma^{-1} \qquad [1]$$

where sums are taken over all selected candidates and k_i will be zero (i.e., unselected) or linearly related to A_i (i.e., selected).

The expression for gain of Bondesson (1989) (i.e., $\sum k_i A_i$) is equivalent to the expression of genetic gain in the context of long-term contributions if clonal proportions are replaced by the long-term genetic contributions and genetic values are replaced by the Mendelian sampling terms (Woolliams and Thompson, 1994). Grundy *et al.* (1998a) showed that the constraint in diversity (γ, in expression [1]) can be expressed in terms of the sum of squares of the long-term contributions, hence in terms of the rate of inbreeding, since $E[\Delta F] = \frac{1}{4}\sum r^2$ (Wray and Thompson, 1990; Woolliams and Thompson, 1994). Hence, the problem can be formulated as:

$$\text{Maximise } \sum r_i a_i \text{ with a constraint } \sum r_i^2 \leq 4T\Delta F \qquad [2]$$

where sums are taken over all male and female candidates and T is the total number of candidates of both sexes per generation ($T/2$ males and $T/2$ females) and ΔF is the desired rate of inbreeding. Bondesson (1989) applied the condition of $\gamma \leq 1$, and in the formulation above this is equivalent to $4T\Delta F \geq 1$. This restriction is satisfied since with a minimum loss of variation, all candidates would be selected and the same contribution would be allocated to each one of them, hence the minimum rate of inbreeding will be $\Delta F \geq \frac{1}{4}\sum \frac{1}{T^2}$, thus showing that $4T\Delta F \geq 1$.

Assuming that the Mendelian sampling terms are normally distributed with standard deviation equal to one, the ideal theoretical rate of genetic gain (ΔG_{ideal}) of Bondesson (1989) can be defined in terms of the standarised truncation point (x) and the selection intensity (i, i.e., the mean deviation of individuals with values above the truncation point):

$$\Delta G_{ideal} = (i - x)^{-1} \text{ or equivalently } i/k \qquad [3]$$

where $k = i(i\text{-}x)$ i.e., the variance reduction factor, and i and x are the solution of

$$(4T\Delta F)^{-1} = p(i - x)^2(1 + x^2 - ix)^{-1}, \qquad [4]$$

where p is the proportion of selected candidates. The truncation point x can be found using the Newton-Rapson method for defined values of $T\Delta F$ (see Appendix 4.1).

The variance of the Mendelian sampling terms is not reduced by selection through gametic phase disequilibrium (i.e., the Bulmer effect, Bulmer, 1971). Hence, the ΔG_{ideal} from expression [3] can be conveniently expressed in terms of the trait additive genetic variance in the base population ($\sigma^2_{A_0}$), resources available (T) and the desired ΔF as:

$$\Delta G_{ideal} = i(k)^{-1}\sqrt{\tfrac{1}{2}\sigma^2_{A_0}}, \qquad [5]$$

where i and x are obtained from [4].

Therefore, for a given $\sigma^2_{A_0}$, number of candidates per generation (T) and target ΔF, the rate of genetic gain that would be obtained after an exact allocation of long-term contributions to the Mendelian sampling terms can be deterministically obtained. This assumes that breeding values and Mendelian sampling terms are known without error (i.e., $h^2_0 = 1.0$ and the variance of environmental effects equal to zero).

It is a valid question to ask what would be the case under the same circumstances of known breeding values and Mendelian sampling terms for truncation selection. That is, how the expression ΔG_{ideal} for an exact allocation of long-term contributions to Mendelian sampling terms compares to conventional formulae for predicting the rate of gain under standard truncation selection. Assuming the phenotypic variance (σ^2_P) equal to one, the expression for the equilibrium rate of gain after accounting for gametic phase disequilibrium (Bulmer, 1971) generated by selection for mass truncation selection can be expressed as:

$$\Delta G_{m\_tru} = \frac{i\sigma^2_{A_0}\sigma^{-1}_P}{(1+kh^2)} \qquad [6]$$

where h^2 is the equilibrium heritability (i.e., h^2 = 1.0 since breeding values are assumed known without error). With the assumption of a base heritability (h^2_0) equal to one, hence the accuracy is also of one, $\sigma^2_{A_0}\sigma^{-1}_P$ is equal to $\sigma^2_{A_0}\sigma^{-1}_A$ where after selection $(1+k)\sigma^2_A = \sigma^2_{A_0}$ (Bulmer, 1971), and $\sigma_A/\sigma_{A_0} = \sqrt{1/(1+k)}$, thus $\sigma^{-1}_A = \sqrt{(1+k)}/\sigma_{A_0}$ and expression [6] can be re-written as:

$$\Delta G_{m\_tru} = \frac{i\sigma_{A_0}}{\sqrt{1+k}} \qquad [7]$$

Equation [7] shows that the form of ΔG_{ideal} in equation [5] is not unexpected and that both expressions for the rate of genetic gain after linear allocation of Mendelian sampling terms and long-term contributions for constrained ΔF, and for truncation selection are closely related. Importantly, it arises that for the same σ_{A_0}, $\Delta G_{m\_tru} = \frac{i}{\sqrt{1+k}} < \Delta G_{ideal} = \sqrt{\frac{1}{2}}\frac{i}{k}$. The validity of this inequality is shown in Figure 4.1, where both expressions for the rate of gain are compared at the *same x and k* (i.e., not at the same ΔF). The ΔG_{ideal} from Grundy *et al.* (1998a), assuming an exact allocation of a and r, and that a are known for each selected candidate (i.e., $h_0^2 = 1.0$), is greater than ΔG_{m_tru} for mass truncation selection when compared at the same truncation point. As expected, the difference was reduced as x and k increased, approaching zero for the limit k=1 (i.e., $\Delta G_{ideal} \cong \Delta G_{m\_tru}$).

As pointed by Grundy *et al.* (1998a) and Woolliams *et al.* (2002), the ΔG_{ideal} is unattainable due to two factors: i) the true a are unknown, so contributions of selected candidates are allocated according to initial estimates of a at the time of selection, and ii) in a multi-generational context, is not possible to manage r of individuals independently. Woolliams *et al.* (2002), stressed out that changing the contribution of an individual alters the contribution of all its ancestors and changing the contribution of a selected parent through its offspring alters the contribution of at least one of its mates.

Grundy *et al.* (1998a) suggested an intuitive expression for predicting the upper bound genetic gain obtained under quadratic optimisation (ΔG_{quad}) that accounts for the lack of knowledge of the true Mendelian sampling terms:

$$\Delta G_{quad} = \rho \times \Delta G_{ideal} , \qquad [8]$$

where ρ is the accuracy of the Mendelian sampling term. Thus, a prediction for ρ is necessary to obtain deterministic predictions for ΔG_{quad}.

4.2.3. Deterministic approach for predicting the accuracy of the Mendelian sampling term

Predictions for the accuracy of the Mendelian sampling term (ρ) were obtained by using the framework of selection index theory. The so-called pseudo-BLUP index of Wray and Hill (1989) was set up to estimate the breeding value of an individual (A_i) under a nested mating structure in which m dams were mated to each sire, and each dam gave n offspring. The total phenotypic variance in generation t is $\sigma^2_{Pt} = \sigma^2_{At} + \sigma^2_C + \sigma^2_E$ where σ^2_{At} is the total additive genetic variation, σ^2_C is the variance of full-sibs common environmental effects and σ^2_E is the environmental variance. In this study, no effects of common environmental effects were assumed (i.e., $\sigma^2_C = 0$). Since the pseudo-BLUP index uses half-sib records, the total additive genetic variance is partitioned as $\sigma^2_{At} = \sigma^2_{ASt} + \sigma^2_{ADt} + \sigma^2_{AWt}$ where σ^2_{ASt} and σ^2_{DSt} are the between family variance for sires and dams, respectively and σ^2_{Awt} is the within family variance (i.e., $\sigma^2_{Awt} = \frac{1}{2}\sigma^2_{A_0}$ were $\sigma^2_{A_0}$ is the additive genetic variance in the base population, before selection). The between family components of the additive genetic variance were updated to account for the effects of selection using the expressions of Wray and Hill (1989).

The original pseudo-BLUP index of Wray and Hill (1989) was extended to accommodate three extra sources of information related to the Mendelian sampling term. This extension allows the accuracy of the Mendelian sampling term to be predicted both when candidates are selected and at the time when they have recorded offspring. The latter accuracy was assumed to be the one at the convergence of the long-term contributions of selected candidates (ρ_{conv}). Expressions for ρ_{conv} were derived for both selected male and female candidates that become sires and dams, respectively

The original pseudo-BLUP index included records of six sources of information: the candidate's own record (X_i), the mean of n (including its own) full-sib records

($\overline{X}_{fs}$), the mean of mn (including the n full–sibs) half sib records ($\overline{X}_{hs}$), the estimated breeding value (EBV) of the dam ($\hat{A}_d$), the mean EBV of all the dams mated to the sire ($\hat{\overline{A}}_d$) and the EBV of the sire ($\hat{A}_s$). The three extra sources of information were: the estimated Mendelian sampling term of the dam ($\hat{a}_d$), the mean estimated Mendelian sampling term of all dams mated to the sire ($\hat{\overline{a}}_d$), and the estimated Mendelian sampling term of the sire ($\hat{a}_s$). Thus the extended index was:

$$I = b_1 X_i + b_2 \overline{X}_{FS} + b_3 \overline{X}_{HS} + b_4 \hat{A}_d + b_5 \hat{\overline{A}}_d + b_6 \hat{A}s + b_7 \hat{a}_d + b_8 \hat{\overline{a}}_d + b_9 \hat{a}_s$$

Index weights were calculated each generation t using standard selection index theory:

$$\mathbf{b}_t = \mathbf{P}_t^{-1} \mathbf{g}_{At},$$

where **b** is the vector of index weights, **P** is the phenotypic variance-covariance matrix for the nine sources of information and $\mathbf{g}_A$ is genetic covariance vector between the breeding value of the individual (A_i) and the nine sources of information. It should be noted that since the new sources of information do not add any information to the estimation of A_i (i.e., the breeding goal), their corresponding index weights (i.e., b_7 to b_9) are zero and the weights for the original sources of information (i.e., b_1 to b_6) remain unchanged.

The accuracy of the Mendelian sampling term was predicted at the time of selection and at the convergence of the long-term contributions of those selected male and female candidates that become sires and dams, respectively. Hence, three extra genetic covariance vectors were calculated between the nine sources of information and the Mendelian sampling term of an individual (a_i). Firstly, at selection time for all candidates ($\mathbf{g}_{a\_cand}$), and secondly after selection when selected male and female candidates become sires ($\mathbf{g}_{a\_sir}$) or dams ($\mathbf{g}_{a\_dam}$), respectively. The **P** matrix and the four g vectors are presented in Appendix 4.2.

Repeated cycles of selection on the extended pseudo-BLUP index were carried out to allow the genetic and phenotypic variance-covariance matrices to reach asymptotic values. The **P** and **g** matrices were updated each generation to account for changes in co(variances) due to selection (Bulmer, 1971). The variance of the estimated breeding values for sires $[V(\hat{A}_s)]$ and dams $[V(\hat{A}_d)]$, that is, the between family components, were updated each generation as described by Wray and Hill (1989).

The newly created covariances resulting from extending the pseudo-BLUP index that required updates in the **P** matrix were the covariances between the estimated breeding value of the individual and the estimated Mendelian sampling term of its sire $[Cov(\hat{A}_s, \hat{a}_s)]$ and its dam $[Cov(\hat{A}_d, \hat{a}_d)]$, and the variance of the estimated Mendelian sampling term of sires $[V(\hat{a}_s)]$ and dams $[V(\hat{a}_d)]$. Likewise, the new genetic covariances created after extending the $\mathbf{g}_A$ vector that were updated each generation were the covariances between the estimated breeding value and the true Mendelian sampling of sires $[Cov(\hat{A}_s, a_s)]$ and the dams $[Cov(\hat{A}_d, a_d)]$. These two latter covariances are actually equal to $Cov(\hat{A}_s, \hat{a}_s)$ and $Cov(\hat{A}_d, \hat{a}_d)$ previously defined, thus their updates were used. The newly defined $\mathbf{g}_{a\_sir}$ and $\mathbf{g}_{a\_dam}$ vectors required the calculation and updating of the covariances between the true Mendelian sampling and the true breeding value of sires $[Cov(A_s, a_s)]$ and the dams $[Cov(A_d, a_d)]$. The complete list of expressions used to obtain asymptotic values of all the variances and covariances in **P** and **g** vectors in the extended pseudo-BLUP index are listed in Appendix 4.3.

Convergence of parameters and index weights was obtained at about generation eight. For all the studied scenarios, equal number of males (n_s) and females (n_d) were selected, hence the mating ratio (d) was equal to one. With d=1, the **P** matrix is not of full-rank as (co)variances involving full- and half-sib terms are equal. Hence, to avoid problems of inversion, those rows and columns in the **P** matrix involving half-sib terms were set to zero, and the corresponding diagonals to one.

4.2.3.1. The accuracy of the Mendelian sampling term

Expressions for the accuracy of the Mendelian sampling term were obtained assuming truncation selection on I. The accuracy of the Mendelian sampling term was calculated for selected candidates just after being selected (ρ_{sel}), and after becoming sires (ρ_{sir}) or dams (ρ_{dam}) with recorded offspring (i.e., assumed as the accuracy at the convergence of their long-term contributions):

$$\rho_{sel} = \sqrt{\frac{V(\hat{a}) - k_{av}[Cov(\hat{A},\hat{a})]^2 / V(I)}{V(a) - k_{av}[Cov(\hat{A},\hat{a})]^2 / V(I)}}, \qquad [9]$$

$$\rho_{sir} = \sqrt{\frac{V(\hat{a}_{sir})}{V(a) - k_s[Cov(\hat{A},\hat{a})]^2 / V(I)}} \quad \text{and} \qquad [10]$$

$$\rho_{dam} = \sqrt{\frac{V(\hat{a}_{dam})}{V(a) - k_d[Cov(\hat{A},\hat{a})]^2 / V(I)}} \qquad [11]$$

where $V(a) = \frac{1}{2}\sigma^2_{A0}$ is the variance of the true Mendelian sampling term, $V(\hat{a}) = \mathbf{g}'_{a\_cand}\mathbf{P}^{-1}\mathbf{g}_{a\_cand}$ is the variance of the estimated Mendelian sampling terms, both in the base population, $Cov(\hat{A},\hat{a}) = \mathbf{g}'_A\mathbf{P}^{-1}\mathbf{g}_{a\_cand}$ is the covariance between the estimated breeding value (i.e., the index) and the estimated Mendelian sampling term of the candidate, $V(I) = \mathbf{g}'_A\mathbf{P}^{-1}\mathbf{g}_A$ is the variance of the index, and k_s, k_d, k_{av}, are the variance reduction coefficients for sires, dams and their average, respectively (i.e., $k = i(i - x)$ with the corresponding i selection intensities and standarised deviations of the truncation point from the corresponding mean). Note that since $k_s = k_d$, it follows that $\rho_{sir} = \rho_{dam} = \rho_{conv}$. The terms $V(\hat{a}_{sire})$ and $V(\hat{a}_{dam})$ correspond to the estimated Mendelian sampling term of selected male and female candidates at the convergence of their long-term genetic contributions. Both variances were calculated

as $V(\hat{a}_{sire}) = \mathbf{g}'_{a\_sir}\mathbf{P}^{-1}\mathbf{g}_{a\_sir}$ and $V(\hat{a}_{dam}) = \mathbf{g}'_{a\_dam}\mathbf{P}^{-1}\mathbf{g}_{a\_dam}$ using asymptotic values for the covariances in $\mathbf{P}$, $\mathbf{g}_{a\_sir}$ and $\mathbf{g}_{a\_dam}$.

The divisor terms in expressions [9], [10] and [11] reflect the fact that after selection, the Mendelian sampling variance of the group of sires and dams is lower than $\frac{1}{2}\sigma^2_{A0}$. The a of a candidate can be written as $\hat{a}+\varepsilon$, where $\hat{a}$ is $b_{\hat{a},I} \times I$ being $b_{\hat{a},I} = \dfrac{Cov(\hat{a},I)}{V(I)}$ the regression of the estimated Mendelian sampling term on the index and I the index value, and ε the prediction error. Accordingly, the Mendelian sampling term variance is $V(a) = V(\hat{a}) + V(\varepsilon)$. After selection $V(a)^* = V(\hat{a})^* + V(\varepsilon)$, where $V(\hat{a})^* = V(\hat{a}) - kCov(\hat{A},\hat{a})^2 / V(I)$ and $V(\varepsilon) = V(a) - V(\hat{a})$. Hence, the remaining Mendelian sampling term variance in the group of selected males and females (i.e., using the corresponding k_s and k_d) is $V(a)^* = \frac{1}{2}\sigma^2_{Ao} - kCov(\hat{A},\hat{a})^2 / V(I)$.

4.2.4. Simulations

The ΔG_{quad} predicted from equation [8] was compared to the empirical responses obtained from stochastic simulations (ΔG_{sim}) over several generations. An additive infinitesimal model (Bulmer, 1971) was considered for the trait under selection. The true breeding values for animals on the base population were obtained from a normal distribution with mean zero and variance equal to the heritability (h^2_0) so the phenotypic variance (σ^2_P) was assumed equal to one. In subsequent generations, the true breeding value of the progeny was obtained as half the sum of the true breeding values of their parents plus a random Mendelian sampling term. The Mendelian sampling term of an individual was sampled from a normal distribution with mean zero and variance $\frac{1}{2}h^2_0$ (i.e., the effects of inbreeding on the additive genetic variance were ignored). The phenotypic value for any individual in any generation was obtained by adding to the true breeding value, an environmental component sampled

from a normal distribution with mean zero and variance $1-h_0^2$. The genetic evaluation to estimate breeding values was carried out using a BLUP animal model. Populations with discrete generations were evaluated over 10 generations of selection. In the base generation (t=0), T individuals ($T/2$ males and $T/2$ females) with family full-sib structure were generated. The first generation of selection was obtained from mating of animals selected at t=0. Populations of T=100, 200 or 300 candidates per generation were simulated per set of parameters. A broad range of trait h_0^2 (0.01, 0.25, 0.50, 0.75 and 0.99) was studied. A total of 100 replicates were performed.

4.2.4.1. Optimised selection with constrained inbreeding

Selection decisions were optimised each generation by using dynamic optimisation tools. The optimisation algorithm described by Meuwissen (1997) for obtaining maximum genetic gain while constraining the inbreeding rate to a specific value was used. This procedure relies on the maximisation of an objective function using Lagrangian multipliers to achieve a restriction in the rate of inbreeding and on the maximum contribution per sex (i.e., one half). The constant rate of inbreeding in the long-term was achieved by setting the constraint on the average coancestry of selected candidates ($\mathbf{c}^{\mathrm{T}}\mathbf{A}\mathbf{c}/2$) to $1-(1-\Delta F)^t$ were $\mathbf{A}$ is the average relationship matrix, $\mathbf{c}$ is a vector of projected use of candidates (i.e., mating proportions), and t is the generation number (Grundy *et al.*, 1998a). With this optimisation procedure, the number of selected candidates and their mating proportions or contributions to the next generation are optimised. The output is a vector of mating proportions ($\mathbf{c}$) of candidates at any particular generation. The optimal number of offspring for an individual i is $2Tc_i$ (a real number), and the actual (integer) number of offspring per parent was obtained following Grundy *et al.* (1998a). Selected candidates were those with $c_i > 0$ and they contributed to the next generation according their c_i values. Mating among selected candidates was at random. The ΔF was constrained to 0.05, 0.025, 0.0125 and 0.01 per generation.

4.2.4.2. *Empirical accuracy of the Mendelian sampling at selection time and at convergence of the long-term contributions*

The predicted accuracy of the Mendelian sampling term from the extended pseudo-BLUP index was compared to empirical accuracy obtained from simulations. The empirical accuracy was calculated as the correlation between the true a and its estimate at two points in time, at selection (ρ_{sel_emp}) and at convergence of the long-term contributions (ρ_{conv_emp}). The estimated Mendelian sampling term for selection candidate i was obtained as $\hat{a}_i = \mathrm{EBV}_i - [\frac{1}{2}(\mathrm{EBV}_S + \mathrm{EBV}_D)]$, where EBV_i, EBV_S, EBV_D are the BLUP estimated breeding values for the individual, its sire and its dam, respectively. The EBV used to calculate $\hat{a}_i$ were either those at the time of selection (EBV_{sel} to estimate $\hat{a}_{sel}$) or at the convergence of long-term contributions of selected candidates (i.e., EBV_{conv} to estimate $\hat{a}_{conv}$).

Selection time and convergence were generation 3 and generation 8 respectively, which corresponds to the group of ancestors and descendants required for the calculation of long-term contributions. The long-term contribution of an ancestor i to a descendant j is defined as the proportion of genes of j that are expected to derive by descent from the ancestor i (Wray and Thompson, 1990; Woolliams *et al.*, 1999). The long-term genetic (r) contributions of ancestors born in generation 3 to descendants born in generation 8 were calculated by tracking back the pedigree from descendants to ancestors using the algorithm described by Woolliams and Mäntysaari (1995). Generation 3 was chosen as the Bulmer effect has already taken place and no further reduction in genetic variance was observed. The term 'at convergence' used throughout the paper refers to generation 8.

4.2.5. *Empirical and predicted effective number of parents at the time of selection and at convergence*

The population structure at the time of selection was required for obtaining the predictions of ρ_{sel} and ρ_{conv} from the extended pseudo-BLUP index. The 'effective number of parents' of equal contributions (e.g., Robertson, 1965) was calculated empirically (i.e., from simulations) at selection time as $N_c = \left[\sum c^2\right]^{-1}$ where the sum of squares was taken over all selected candidates with optimised mating proportions. The population structure at selection time was derived as $n_s = n_d = \frac{1}{2}N_c$ and the number of offspring per dam was T/n_d. Also, the 'effective number of ancestors' was calculated from the long-term contributions as $N_r = \left[\sum r^2\right]^{-1}$ where the sum of squares was taken over all male and female selected candidates in generation 3 whose r to generation 8 were calculated.

Since the primary objective was to derive a completely deterministic prediction of ΔG_{quad}, an approach was derived for predicting N_c. The prediction was based on the empirical ratio N_c/N_r which measures the change in the equivalent number of parents from initial selection that are still represented when their contributions converge. The N_r can be derived directly from the ΔF constraint as $N_r = \left[4\Delta F\right]^{-1}$ since $E(\Delta F) = \frac{1}{4}\sum r_i^2$. Therefore, N_c could be obtained from a prediction of N_c/N_r and the ΔF constraint. The prediction approach for the ratio N_c/N_r was derived after following three steps. Firstly, the observed ratios were plotted against $1-h_0^2$ for different values of the joint factor $T\Delta F$ (Figure 4.4.1, Appendix 4.4). The use of $1-h_0^2$ rather than h_0^2 reflects better the effect of the accumulation of information as h_0^2 increases (e.g., for $\sigma_P^2 = 1.0$, $1-h_0^2 = \sigma_e^2$). Also, the factor $T\Delta F$ was chosen since it represents an intrinsic design variable of the breeding scheme, and is convenient from a dimensionality point of view as it allows fitting less number of parameters. Secondly, the regression coefficients of the double natural logarithm (ln) of N_c/N_r on $1-h_0^2$ for several levels of $T\Delta F$, namely $b_{T\Delta F}$ were estimated from a nested regression model such as $\ln(\ln Nc/Nr) = a_1 + b_{T\Delta F}(1-h_0^2)$ [R1], were a_1 is the intercept (Figure 4.4.2, Appendix 4.4). Thirdly, the regression coefficients for each level of $T\Delta F$ were regressed on $\ln(T\Delta F)$ as $b_{T\Delta F} = a_2 + b_2 \ln(T\Delta F)$ [R2]

where a_2 is the intercept and b_2 the regression coefficient (Figure 4.4.3, Appendix 4.4). A full prediction model can be written by substituting [R2] in [R1] as $\ln(\ln N_c/N_r) = a_1 + [a_2 + b_2 \ln(T\Delta F)] \times (1 - h_0^2)$. After taking anti-logarithms, the prediction of N_c/N_r can be expressed in terms of $1 - h_0^2$ and $T\Delta F$ as $\ln(N_c/N_r) = e^{a_1} \times T\Delta F^{b_2} \times e^{a_2 \times (1-h_0^2)}$. The prediction approach is presented in Appendix 4.4 for the range of parameter values used for $1 - h_0^2$ and $T\Delta F$.

4.2.6. Comparison of deterministically predicted genetic gains from optimal selection and truncation selection at the same rate of inbreeding

The ΔG_{quad} predicted from equation [8] was compared to the rate of gain obtained from truncation selection at the same ΔF for a range of breeding schemes. The predictions of genetic gain for truncation schemes were obtained by using the software Sel*A*ction (Bijma and Rutten, 2002; Rutten *et al.*, 2002). Sel*A*ction has been designed for allowing deterministic predictions of both ΔG and ΔF for specific sets of design variables, including genetic and phenotypic parameters and population structure (e.g., number of selected candidates per sex, number of offspring per dam, etc.). The response to selection is predicted based on the pseudo-BLUP index theory (Wray and Hill, 1989) and accounts for the reduction of variance due to selection (Bulmer, 1971). The rate of inbreeding is predicted based upon the theory of long-term genetic contributions (Woolliams and Thompson, 1994) and uses algorithms developed by Woolliams and Bijma (2000), Bijma *et al.* (2000) and Bijma and Woolliams (2000). Details of the theoretical framework that underlies Sel*A*ction are provided in Bijma *et al.* (2001).

Comparisons between truncation and optimal selection at the same rate of inbreeding were carried out in two ways. Firstly, ΔG_{quad} and ΔG_{tru} were compared for an extensive range of ΔF at two fixed levels of h_0^2 and secondly, ΔG_{quad} and ΔG_{tru} were compared for an extensive range of h_0^2 and two fixed levels of ΔF.

For the comparison across ΔF, the difference between ΔG_{quad} and ΔG_{tru} was predicted for h_0^2= 0.10 or 0.35 and T=100 or 300 candidates. For a given T, the population structure was derived using d =1 (i.e., $n_s = n_d$ = T/2 and the number of offspring per dam, n = T/n_d). To obtain an extensive range of ΔF (from 0.1% up to 4.0%) the population structure was varied by changing the number and proportion of selected males and females (i.e., $p_m = p_f$) at regular intervals starting from 0.96. For each population structure, Sel*A*ction predicted ΔG_{tru} and the corresponding ΔF. For each level of ΔF predicted under truncation selection, the N_c was predicted from T, ΔF and h_0^2 using the model described in the previous section. The ΔG_{quad} under optimal selection was calculated after predicting ρ_{conv} and ΔG_{ideal} from the modified pseudo-BLUP index and Equation [5], respectively.

For the comparison across h_0^2 the ratio $\Delta G_{quad} / \Delta G_{tru}$ was predicted for a range of h_0^2 from 0.05 to 0.85 and T=100, 300 and 1000, and two constraints on ΔF of 0.01 and 0.025. For a given T, the population structure that gave the desired ΔF under truncation selection was first found using Sel*A*ction for each level of h_0^2. In most cases the desired ΔF was achieved, although in a few cases a maximum deviation of 0.05 with respect to the target ΔF was allowed. Then, the corresponding N_c was predicted from T, ΔF and h_0^2, and ΔG_{quad} was calculated as described above.

4.3. Results

4.3.1. Empirical optimal population structures

Table 4.2 shows the empirical optimal population structures that maximise ΔG for a given level of ΔF obtained from simulations at selection time and at convergence. The simulation results show that the relationship between the sum of squares of contributions at selection time ($\sum c^2$) and at the convergence of the long-term

contributions ($\sum r^2$), hence between N_c and N_r is primarily a function of the trait h_0^2 and is less dependent on the ΔF constraint (see N_c/N_r values). The ratio N_c/N_r measures the departure from the ideal optimal solution of an exact allocation of long-term contributions to Mendelian sampling terms and thus it can be viewed as a measure of efficiency of the scheme (Grundy *et al.*, 1998a; Woolliams *et al.*, 2002). The ratio N_c/N_r increased as h_0^2 decreased and as the ΔF constraint was less stringent. For instance, for h_0^2=0.01 and ΔF=0.05 the N_c was more than twice N_r (i.e., $N_c/N_r = 2.17$), whereas for h_0^2=0.99 and ΔF=0.01 the agreement between N_c and N_r was much better (i.e., $N_c/N_r = 1.24$).

One important observation is that for a given ΔF constraint N_r was the same across the whole range of h_0^2, and that the corresponding $\frac{1}{4}\sum r^2$ was equal or lower than the target ΔF constraint.

4.3.2. Predicted accuracy of the Mendelian sampling term

The values of N_c showed in Table 4.2 were used to derive the population structure required to obtain the predictions of ρ_{sel} and ρ_{conv} with the modified pseudo-BLUP index using d =1 (i.e., $n_s = n_d = N_c/2$ and the number of offspring per dam, $n = N_c/n_d$).

The relationship between empirical and predicted accuracy of the Mendelian sampling term at selection time and at convergence are shown in Figure 4.2. The extended pseudo-BLUP index allowed for reasonably good predictions of the Mendelian sampling accuracy conditional to N_c. Both at selection and at convergence the regression of the predicted accuracy on the observed accuracy was close to 1.00, and the correlation between them was about 0.99. In particular, the predictions of ρ_{conv}, the critical parameter for predicting ΔG_{quad} were particularly accurate.

4.3.3. Deterministic prediction of the rate of genetic gain under quadratic optimisation

The ΔG_{quad} was predicted from equation [8] for the range of h_0^2, ΔF constraint and population structures presented in Table 4.3 (i.e., for T=100 candidates). The Mendelian sampling term accuracy at convergence was predicted from the extended pseudo-BLUP index for sires and dams (i.e., equations [9] and [10] for ρ_{sir} and ρ_{dam}, respectively) using empirical population structures derived from N_c. Since the mating ratio was set to 1 the ρ_{sir} and ρ_{dam} were equal.

Results presented in Table 4.3 show good agreement between ΔG_{obs} and ΔG_{quad} for the broad range of h_0^2 and ΔF constraints studied. The %error ranged from –0.7% (h_0^2=0.5and ΔF=0.025) up to 15.0% (h^2=0.99 and ΔF=0.01). Over-prediction was the norm for the most relaxed ΔF constraint (0.05), for ΔF = 0.025 at h_0^2 of 0.01 and 0.25, and for all ΔF constraints at h_0^2 = 0.99. On the other hand, under-prediction was typically observed for the most stringent ΔF constraints (0.0125 and 0.01). Nevertheless, very good agreement (e.g., %error less than 10%) between ΔG_{obs} and ΔG_{quad} was observed for h_0^2 levels typically found in production traits (e.g., up to 0.50) and for the range ΔF of constraints most likely to be applied in practice (e.g., from 0.01 to 0.025).

Predictions of ΔG_{quad} for a broad range of h_0^2 are compared with the rates of gain from simulations in Figure 4.3 for two scheme sizes (T=100 and T=300), and two ΔF constraints (ΔF=0.01 and ΔF=0.025). Both predictions of ΔG_{quad} using empirical and predicted N_c are presented. The predictions of ΔG_{quad} obtained after using the predicted N_c from the regression model on h^2 and $T\Delta F$ are as good as using empirical N_c for both T=100 (Figure 4.3.a and 4.3.b) and T=300 (Figure 4.3.c and 4.3.d). Thus, only the differences between ΔG_{quad} and ΔG_{obs} based on the predicted N_c are described here. A very good agreement between ΔG_{quad} and ΔG_{obs} was observed for the smaller scheme (T=100) at both ΔF constraints (i.e., excluding the upper limit

h_0^2=0.99). For ΔF=0.01, only a small underprediction of on average -5.9% was observed, and the differences between predicted and observed values were non-significant (p>0.05) for h_0^2 ranging from 0.01 to 0.7. For ΔF=0.025, overprediction of on average 3.4% up to h_0^2=0.4 and an underprediction of on average –2.4% from h_0^2=0.5 was observed. Nevertheless, the difference between ΔG_{quad} and ΔG_{obs} was non-significant (p>0.05) for h_0^2 up to 0.6.

For the largest scheme (T=300), the ΔG_{quad} still provided good predictions of gain, particularly for the most stringent constraint. For ΔF=0.01, the difference between ΔG_{quad} and ΔG_{obs} was non-significant (p>0.05) up to h_0^2=0.9 (Figure 4.3.b). A small overprediction of about 3.9% was observed up to h_0^2=0.7, and a small underprediction of –3.7% was observed from h_0^2=0.6. For ΔF=0.025, the gain was overestimated on average 11.8% in virtually the whole range of h_0^2 up to 0.8. The difference between ΔG_{quad} and ΔG_{obs} was significant (p<0.05) for h_0^2 ranging from 0.1 to 0.6, and showed a peak of 22.4% for h_0^2=0.2 (Figure 4.3.d).

4.3.4. Comparison of predicted rates of genetic gain from truncation and quadratic optimisation at the same rate of inbreeding

The rates of gain under truncation and optimal selection for a range of ΔF levels are presented in Figure 4.4. The ΔG_{quad} was higher than ΔG_{tru} at all levels of ΔF for both trait heritabilities and scheme sizes. The advantage of ΔG_{quad} over ΔG_{tru} increased as the ΔF increased (i.e., a more relaxed constraint in quadratic optimisation) and was greatest for the highest h_0^2 and the largest breeding scheme. For instance, for ΔF=0.01 and h_0^2=0.35, the advantage of ΔG_{quad} over ΔG_{tru} was 26.5% (0.286 vs. 0.226, respectively) for T=100 and 32.5% (0.473 vs. 0.357, respectively) for T=300. For ΔF=0.01 and h_0^2=0.10 the advantage of ΔG_{quad} over ΔG_{tru} was 17.3% (0.095 vs. 0.081, respectively) for T=100 and 30.4% (0.163 vs. 0.125, respectively) for T=300.

The same profile for the advantage of ΔG_{quad} over ΔG_{tru} presented in Figure 4.4 was found when T was increased up to 1,000 candidates. In this scenario, the superiority of ΔG_{quad} over ΔG_{tru} was about 40% at ΔF=0.01 (results not shown).

Figure 4.5 shows the ratio $\Delta G_{quad} / \Delta G_{tru}$ for a range of h_0^2 and three population sizes at fixed values of ΔF of 0.01 and 0.025. Two situations can be identified according to ranges of h_0^2 below or above 0.55. Firstly, for h_0^2 up to 0.55, the ratios $\Delta G_{quad} / \Delta G_{tru}$ were greater for ΔF=0.025 than for ΔF=0.01, and a maximum at around h_0^2= 0.15 or 0.25 was observed for T=300 and T=1,000 at ΔF=0.01, and for the three population sizes at ΔF=0.025. The maximum ratio increased with the population size. For ΔF=0.01, the maximum $\Delta G_{quad} / \Delta G_{tru}$ were 1.41 for T=1000, 1.32 for T=300 and 1.27 for T=100, at h_0^2=0.15, 0.25 and 0.45, respectively. For ΔF=0.025, the maximum $\Delta G_{quad} / \Delta G_{tru}$ were 1.45 for T=1000, 1.40 for T=300 and 1.32 for T=100, at h_0^2=0.15, 0.25 and 0.25, respectively. For h_0^2 above 0.55, at ΔF=0.01 the ratio $\Delta G_{quad} / \Delta G_{tru}$ was fairly constant up to h_0^2=0.85 for T=100, whereas for T=300 and T=1,000 it decreased down to 1.15 and 1.13, respectively. For ΔF=0.025, the ratio $\Delta G_{quad} / \Delta G_{tru}$ decreased for the three population sizes reaching a minimum value of 1.16, 1.13 and 1.12 for T=100, 300 and 1000, respectively at h_0^2=0.85. This indicates that the effect of the size of the scheme on the superiority of quadratic optimisation over truncation selection decreases as the trait h_0^2 increases (e.g., with the exception of very small schemes at tight ΔF constraints).

It should be noted that the predictions of ΔG_{quad} for T=1,000 were obtained from a N_c prediction that was out of the range of the empirical data used to obtain the prediction equation for N_c described previously. Nevertheless, this seems not to affect the results presented in Figure 4.4, as the ratio $\Delta G_{quad} / \Delta G_{tru}$ followed the same trend across h_0^2 as the corresponding ratios for T=100 and 300, for which the N_c prediction was supported by empirical data.

4.4. Discussion

This study has presented a deterministic approach for the prediction of the optimal rate of genetic gain in schemes in which the rate of inbreeding is restricted to pre-defined levels. The two key components required for the prediction are: i) a prediction of the ideal rate of gain after a perfect allocation of long-term contributions to Mendelian sampling terms for a given ΔF, and ii) a prediction of the ultimate accuracy of the Mendelian sampling term at the time of convergence of the long-term contributions of selected candidates. The first component (i.e., prediction of ΔG_{ideal}) was available from Grundy *et al.* (1998a). Here, an extended pseudo-BLUP index selection has been developed, that allowed for accurate predictions of ρ both at selection time and at convergence. The prediction framework was completed after providing a prediction of the sum of squares of the mating proportions, that is, the optimisation outcome at the time of selection. The present framework provided accurate predictions of ΔG for a broad range of h_0^2, ΔF constraints and breeding schemes of contrasting size.

Deterministic methods for the strategic optimisation of breeding schemes, that is, the maximisation of genetic gain for a pre-defined level of the rate of inbreeding were previously developed for schemes under mass selection with discrete (Villanueva *et al.*, 1996) and overlapping generations (Villanueva *et al.*, 2000), and for schemes under index selection and discrete generations (Villanueva and Woolliams, 1997). However, these methods were developed under truncation selection, and were unlinked to dynamic selection algorithms based upon quadratic indices (Meuwissen, 1997; Grundy *et al.*, 1998a). This prediction framework differs from that of Villanueva *et al.*, (1996); Villanueva and Woolliams, (1997) and Villanueva *et al.*, (2000) not only in the selection method to which is referred (i.e., quadratic optimisation), but also in the type of answers that provides. The existing framework for deterministic optimisation under truncation selection required a prediction of the ΔF and provided the detailed optimal scheme design in terms of optimal number of selected males and females (i.e., the breeding structure) for a set of design variables

(i.e., h_0^2, scheme size and ΔF). In contrast, in this approach, the ΔF is a known variable and the maximum rate of gain is provided for a specific scheme size, but without giving details on a particular breeding structure. It should be emphasised though, that this approach is not an extension of previous deterministic framework for mass selection or sib-index selection to BLUP selection, but it is a completely new approach that deals with quadratic optimisation.

The framework of Villanueva *et al.*, (2000), although for truncation selection, was the first deterministic optimisation that explicitly defined the rate of gain in terms of the long-term contributions and the Mendelian sampling terms (i.e., $E[\Delta G] = \sum r_i a_i$, Woolliams and Thompson, 1994). Both predictions of ΔG and ΔF were based on predictions of the lifetime contributions (u_i) of selected candidates since $E[\Delta G] = \sum E[u_i a_i]$ and $E[\Delta F] = \frac{1}{2}\sum E[u_i^2]$ (Woolliams *et al.*, 1999). The basic framework for predicting u_i was laid down by Woolliams *et al.*, (1999) and is based on $u_i = E(r_i \mid s_i)$ where s_i is the selective advantage of the individual i. The expected long-term contribution of an ancestor i in category q (male or female) was given by $u_i = \alpha_q + \beta_j(s_i - \bar{s}_q)$, where α_q is the expected contribution of an average parent in category q and β_j is the regression of the contribution of the individual i on its selective advantage (expressed as a deviation from selected contemporaries in the same category $\bar{s}_q$). The prediction of u_i was the key parameter for the prediction of ΔF in livestock populations, and formula is available for practical breeding programs under truncation selection, BLUP evaluation and overlapping generations (e.g., Bijma *et al.*, 2001).

Although the framework for predicting lifetime contributions was relevant under truncation selection it may not be appropriate when quadratic indices are used as pointed out by Woolliams *et al.*, (1999). For instance, under truncation selection selective advantages are defined as functions of the candidates breeding value whereas under quadratic optimisation the selective advantage is related to the Mendelian sampling term (see Chapter 3). The main limitation for extending the

prediction approach based on u_i and developed under truncation selection to constrained optimisation under BLUP selection has been identified as the difficulty of predicting the selection intensity for the group of selected candidates for which the average coancestry have been restricted (Bijma, 2000, Thesis General Discussion; Woolliams *et al.*, 2002).

Here, the route of predicting the accuracy of the Mendelian sampling term as the key parameter for obtaining predictions of ΔG under constrained ΔF was followed. The present approach is clearly linked to the ideal solution for the constrained optimisation problem presented by Grundy *et al.*, (1998a) and the empirical findings in Chapter 3, that at all times, from selection to convergence of the long-term contributions, quadratic indices attempt to manage contributions in relation to the best information on the Mendelian sampling term and not the breeding value. The lack of knowledge of the Mendelian sampling term and the availability of only initial estimates at the time of selection is one of the central factors that prevent attaining the ideal optimal solution of an exact allocation of r and a (Grundy *et al.*, 1998a; Woolliams *et al.*, 2002). Therefore, by developing accurate predictions of ultimate ρ the over-prediction represented by ΔG_{ideal} is directly adjusted. The inability to set the r of selected candidates to their desired values (i.e., independently from the contributions of future descendants) is the second factor that drives the observed outcome of the constrained optimisation away from the ideal solution (Grundy *et al.*, 1998a; Woolliams *et al.*, 2002). Not accounting for this effect when adjusting ΔG_{ideal} seems not to represent a strong limitation as accurate predictions for a broad range of parameters have been obtained.

The idea of expressing the Mendelian sampling term as the selective advantage was first put forward by Woolliams and Thompson (1994). They expressed the breeding value of an individual as the weighted sum across generations of the Mendelian sampling terms of its ancestors with the weighing factors being the ancestors long-term contributions (i.e., $A_{j,t} = \sum_{u=1}^{t} \sum_{i} r_{i,u}(j,t) a_{i,u}$, for an individual j born at time t). This idea gave the scope for the formal derivation of $E[\Delta G] = \sum E[r_i a_i]$ provided by

Woolliams *et al.*, (1999). Woolliams and Thompson (1994) also laid down the idea of constructing Mendelian indices by decomposing the estimated breeding value as $\hat{A}_j = \hat{a}_j + \sum_{t=1}^{T}\sum_{i=1}^{2^t} \hat{a}_{ji,t} c^t$, where $\hat{a}_j$ is the estimated Mendelian sampling term of the individual j and $\hat{a}_{ji}$ represents the Mendelian sampling term of the ancestor i born with t generations separating the candidate and the ancestor. The factor c^t represents the weight given to the Mendelian terms of ancestors thus, to the family information. Grundy *et al.,* (1998b) implemented such linear Mendelian indices as a way of managing genetic gain and inbreeding in a flexible way, but although they proved to be successful in managing inbreeding, they typically imposed costs in selection accuracy and in genetic gain in the short term. Furthermore, the generalisation of their effects across discrete or overlapping generations was difficult. With the present approach of predicting the accuracy of the Mendelian sampling term it is recognised the relationship between the selective advantage and genetic gain explicited by Woolliams and Thompson (1994), and it is acknowledged the fact that quadratic optimisation explicitly maximises genetic gain for a pre-defined level of inbreeding. Moreover, the idea of the selective advantage being the Mendelian sampling term, and the implementation of a pseudo-BLUP index including terms related to its estimates agrees with the concept of Mendelian indices of Woolliams and Thompson (1994).

The prediction of ultimate ρ by using the extended pseudo-BLUP assumed convergence after one generation, that is, when selection candidates had recorded offspring available. Although in practice long-term contributions converge after about four to five generations (e.g., Woolliams and Mäntysaari, 1995), the assumption of convergence is valid as the r of an individual will be mainly affected by its initial contribution (i.e., the mating proportions) and the contribution of its offspring. In addition, ρ converges rapidly as the biggest impact comes from the offspring information whereas the grand-offspring and grand-grand offspring information will have only a marginal impact. Therefore, by assuming the ultimate ρ one generation after selection, it will coincide with the time at which ρ will be near convergence and when most of the impact in changing r would have been made as

well. The agreement between observed and predicted ρ at the convergence of the long-term genetic contributions showed in Figure 4.2 reinforces the validity of this assumption. A further simplification of these predictions of ρ is the assumption of equal usage of selected candidates (i.e., specified in the pseudo-BLUP index), thus the unequal contributions that characterises quadratic indices are not directly accounted for. Nevertheless, the predictions of ρ are derived after predicting the sum of squares of contributions at the time of selection that would have been obtained after applying quadratic optimisation for a desired ΔF constraint, thus directly accounting for that selection method.

Although the deterministic framework for predicting genetic gain is self-contained and can be used by the specification of solely T, h_0^2 and ΔF, it should be recognised that there is an element of empirical nature in the approach. This arises from the regression approach for predicting the ratio N_c/N_r that used the empirical relationship between the optimisation outcome at selection time ($\sum c^2$) and at convergence ($\sum r^2$) for different h_0^2, ΔF restriction and population sizes. Nevertheless, the prediction of N_c/N_r should not be seen as a limitation since it not only provides an intuitive way of predicting N_c, but has also highlighted the need for a deterministic framework for describing the process of moving from selection to convergence under constrained optimisation using quadratic indices. During the development work it was found that the current approach for predicting the ratio N_c/N_r is robust to changes in the model used for representing the empirical relationship between N_c/N_r and T, h_0^2, and ΔF (e.g., linear or quadratic regressions).

The present framework ignored the effects of inbreeding on the genetic variance, thus predictions refer to an asymptotic response to selection only affected by gamete phase disequilibrium (i.e., the Bulmer effect). Villanueva and Woolliams (1997) found with sib-indices that under restricted inbreeding the optimum schemes for maximising gain at generation 5 or 20 were similar when the effect of inbreeding on

genetic variance was accounted for. In essence, by applying a restriction in the amount of inbreeding, the loss of genetic variance is also restricted. In practice, responses in populations under artificial selection have been maintained in the long-term, thus suggesting that the effects of effective population size on genetic variance might not be well described under the infinitesimal model (see Hill and Zhang, 2003 for a review). Thus, ignoring the effects of inbreeding on genetic variance might not represent a limitation from a modelling viewpoint. In any case, here, the predictions of the rate of gain were compared to the rate of gain obtained from simulations without including reduction of genetic variance due to inbreeding, thus not including sources of bias.

A relevant result of this study is the deterministic comparison of the ΔG for schemes under optimal and truncation selection at the same ΔF. This is the first time such comparison is available under BLUP selection and provides clear comparison of the relationship between ΔG and ΔF for both selection methods. Results indicate that for any level of ΔF a higher ΔG is predicted under optimal selection, and that for h_0^2 up to 0.55 the difference will increase as the population size increases. Results in Figure 4.5 suggest that the maximum advantage of quadratic optimisation (ΔG_{quad}) over truncation selection (ΔG_{tru}) occurs at h_0^2 ranging from 0.15 to 0.35, for ΔF values of 0.01 and 0.025 which represent likely values for a target rate of inbreeding in commercial breeding populations. In addition, for traits with h_0^2 ranging from 0.05 to 0.55 (e.g., from reproductive, fitness or disease related traits to carcass related traits) the superiority of ΔG_{quad} over ΔG_{tru} will be greater as the size of the scheme increases. Benchmark maximum values for the ratio $\Delta G_{quad} / \Delta G_{tru}$ of 1.40 and 1.45 for the largest scheme (i.e., T=1,000), and 1.27 and 1.32 for the smallest scheme (i.e., T=100) for ΔF values of 0.01 and 0.025, respectively were found. The predicted superiority of quadratic optimisation over truncation selection agree with empirical evidence found in real livestock populations of Aberdeen Angus (beef cattle) and Meatlinc (sheep) in Chapter 2. At the observed ΔF the expected ratio $\Delta G_{quad} / \Delta G_{tru}$

was 1.30 and 1.17 for Aberdeen Angus (ΔF = 0.002) and Meatlinc (ΔF = 0.01), respectively.

Three points should be addressed in relation to the predicted superiority of ΔG_{quad} over ΔG_{tru}. First, on the increasing $\Delta G_{quad} / \Delta G_{tru}$ with greater scheme sizes. With larger schemes, larger family sizes are created having an impact on both the selection intensity and the accuracy of the Mendelian sampling term. In a larger scheme, the required selection intensity for meeting the ΔF constraint can be more easily achieved, hence, the selection intensity on the selective advantage (i.e., the Mendelian samping term under quadratic optimisation) is not compromised, in contrast it is expected to be greater than in smaller schemes. In contrast, in smaller schemes, the smaller family sizes impose more constraints in the use of candidates to achieve the desired ΔF and to maximise ΔG. In addition, the accuracy of the Mendelian sampling term is greater in larger schemes. For instance, for h_0^2 = 0.35 and ΔF = 0.01 the predicted ρ was 0.67 for T = 1000, 0.57 for T = 300 and 0.48 for T = 100 (i.e., the same was observed from simulations for T = 300 and 100). Second, a different family structure might be contributing to the superiority of ΔG_{quad} over ΔG_{tru}. When the mating ratio is one, only full-sib families are created under truncation selection, whereas under quadratic optimisation the way in which the random mating is implemented enables the creation of maternal half-sib families (e.g., in a factorial mating fashion). Under quadratic optimisation a female with an optimal c_i equivalent to more than one mating will be allocated more than one male creating maternal half-sib families. Nevertheless the effect of the family structure on $\Delta G_{quad} / \Delta G_{tru}$ is expected to be lower as the scheme size increases (e.g., Sonesson and Meuwissen, 2000). Third, from a theoretical point of view, the ratio $\Delta G_{quad} / \Delta G_{tru}$ might be somewhat overestimated since truncation selection with BLUP selection is not the best achievable truncation selection strategy. Sel*A*ction predicts the genetic gain under truncation selection that would be obtained in practical breeding schemes using BLUP, but even higher gains might be achieved if the weights given to the six sources of information in the pseudo-BLUP index were optimised. Villanueva and Woolliams (1997) found benefits from optimising the

relative weights in a family index with three sources of information (i.e., individual records, half-sib and full-sib group mean records) when maximising ΔG for a given ΔF under truncation selection. They envisaged that the optimisation of index weights procedure could be extended to the pseudo-BLUP, although the implementation could be complex. Hence, if optimal relative weights for the pseudo-BLUP index were used, the ΔG_{tru} might be higher and the $\Delta G_{quad} / \Delta G_{tru}$ reduced.

These results are critical from the point of view of the strategic design of practical breeding schemes, since with this framework breeders will be able set their risk preference (i.e., the ΔF) and predict the benefit from changing from their current selection practices (namely, truncation selection) to optimal selection. In a subsequent stage, available operational tools (e.g., Meuwissen, 1997; Grundy *et al.*, 1998a) for running the optimisation of the breeding scheme on a day-to day basis can be applied. The key element is that now an '*a priori*' (i.e., deterministic) optimisation that takes account of the selection index used (i.e., BLUP) is available. Moreover, both the design tool and the '*a posteriori*' (i.e., operational) tool have the same underlying definition of genetic gain expressed in terms of long-term genetic contributions and Mendelian sampling terms of Woolliams and Thompson (1994). Thus, the prediction framework for ΔG provides the necessary accompanying tool for the available operational dynamic optimisation algorithms.

As it stands today, the prediction framework can be readily applied in species with no restriction in reproductive rates (e.g., fish) to improve traits with phenotypes available in both sexes. Extensions of this prediction framework, particularly to accommodate reproductive limitations would be relevant to allow its application in cattle and sheep populations. The current extended pseudo-BLUP index deals with any mating ratio and thus predictions of the Mendelian sampling accuracy are not a limitation. In contrast, an expression for the ideal optimal ΔG when reproductive limitations exist needs to be developed.

This study has filled the gap in prediction framework by providing for the first time a straightforward way of solving $E(\Delta G \mid \Delta F)$ by only specification of the amount of

resources (summarised by T), the trait h_0^2, the target ΔF and the corresponding sum of squares of mating proportions (i.e., N_c).

4.5. Tables

Table 4.1. Main notational conventions

ΔF	Rate of inbreeding
ΔG_{ideal}	Predicted ideal rate of genetic gain under constrained inbreeding
ΔG_{m_tru}	Predicted rate of genetic gain under mass truncation selection
ΔG_{tru}	Predicted rate of genetic gain under BLUP truncation selection
ΔG_{quad}	Predicted ate of genetic gain under quadratic optimisation
ΔG_{obs}	Observed rate of genetic gain from simulations
c_i	Contribution to the next generation of individual i
r_i	Long-term contribution of individual i
A_i	Breeding value of individual i
EBV_{sel}, EBV_{conv}	Estimated breeding values of selected candidates at selection time and at the convergence of the long-term contributions
$\hat{A}_s$, $\hat{A}_d$	Estimated breeding value of sires and dams
a_i	Mendelian sampling term of individual i
$\hat{a}_{sel}$, $\hat{a}_{conv}$	Estimated Mendelian sampling term at selection time and at convergence of the long-term contributions
$\hat{a}_s$, $\hat{a}_d$	Estimated Mendelian sampling term of sires and dams
$\sigma^2_{A_0}$, σ^2_A	Trait additive genetic variation in the unselected base population and after selection
σ^2_P	Trait phenotypic variance
h^2_0, h^2	Trait heritability in the unselected base population and after selection
ρ_{sel}, ρ_{conv}	Predicted accuracy of the Mendelian sampling term at selection and at convergence of the long-term contributions of selected male and female candidates
ρ_{sir}, ρ_{dam}	Predicted accuracy of the Mendelian sampling term of male and female parents at the convergence of their long-term contributions

T	Number of candidates of both sexes per generation
d	Mating ratio, number of females mated to a single sire
n, n_s, n_d	Number of offspring per dam, number of male and female parents
N_c, N_r	Effective number of parents of both sexes at selection time and at convergence of the long-term contributions
$\mathbf{P}$	Phenotypic varianc-covariance matrix for sources of information in the extended pseudo-BLUP index
$\mathbf{g}_A$	Genetic covariance vector between the breeding value and the sources of information in the extended pseudo-BLUP index
$\mathbf{g}_{a\_cand}$	Genetic covariance vector between the Mendelian sampling term of all candidates at selection time and the sources of information in the extended pseudo-BLUP index
$\mathbf{g}_{a\_sir}$, $\mathbf{g}_{a\_dam}$	Genetic covariance vector between the Mendelian sampling term of sires and dams at convergence of the long-term contributions the sources of information in the extended pseudo-BLUP index
$V(X)$	Variance of variable X
$Cov(X,Y)$	Covariance between variables X and Y

Table 4.2. Empirical sum of squares of contributions and effective number of sires at selection time ($\sum c^2$ and N_c, respectively) and at convergence of long-term contributions ($\sum r^2$ and N_r, respectively) and the ratio N_c/N_r obtained from optimal selection for a range of heritabilities (h_0^2) and constraints in the rate of inbreeding (ΔF), and a population size of 100 candidates per generation.

		Selection time (t=3)		Convergence (t=8)		
h_0^2	ΔF	$\sum c^2$	N_c^{-1}	$\sum r^2$	N_r	N_c/N_r
0.01	0.05	0.0795	13	0.1680	6	2.18
	0.025	0.0469	21	0.0910	11	1.91
	0.0125	0.0271	37	0.0459	22	1.70
	0.01	0.0231	43	0.0375	27	1.61
0.25	0.05	0.0996	10	0.1720	6	1.72
	0.025	0.0552	18	0.0922	11	1.66
	0.0125	0.0307	33	0.0476	21	1.57
	0.01	0.0255	39	0.0375	27	1.46
0.5	0.05	0.1122	9	0.1684	6	1.52
	0.025	0.0607	16	0.0909	11	1.45
	0.0125	0.0334	30	0.0470	21	1.41
	0.01	0.0275	36	0.0376	27	1.35
0.75	0.05	0.1274	8	0.1693	6	1.35
	0.025	0.0674	15	0.0906	11	1.36
	0.0125	0.0354	28	0.0472	21	1.32
	0.01	0.0290	34	0.0375	27	1.28
0.99	0.05	0.1405	7	0.1740	6	1.22
	0.025	0.0717	14	0.0905	11	1.27
	0.0125	0.0374	27	0.0470	21	1.27
	0.01	0.0302	33	0.0375	27	1.24

[1] $N_c = \left(\sum c^2\right)^{-1}$ and $N_r = \left(\sum r^2\right)^{-1}$

Table 4.3. Rates of gain obtained from simulations[1] (ΔG_{obs}), theoretical ideal upper limit (ΔG_{ideal}), the predicted rate of gain from quadratic optimisation (ΔG_{quad}), and predicted accuracy of the Mendelian sampling term for sires (ρ_{ult}) at the time of convergence of the long-term contributions for a range of h_o^2 and ΔF constraints. The population size was 100 candidates per generation.

h_o^2	ΔF	ΔG_{obs}[2]	ΔG_{ideal}[3]	ρ_{ult}[4]	ΔG_{quad}[5]	% error[6]
0.01	0.05	0.021	0.153	0.148	0.023	7.8
	0.025	0.015	0.132	0.123	0.016	8.0
	0.0125	0.011	0.108	0.100	0.011	-3.5
	0.01	0.010	0.099	0.096	0.009	-4.4
0.25	0.05	0.374	0.766	0.558	0.427	14.3
	0.025	0.316	0.660	0.495	0.327	3.4
	0.0125	0.250	0.538	0.435	0.234	-6.5
	0.01	0.227	0.494	0.419	0.207	-9.0
0.5	0.05	0.668	1.083	0.673	0.729	9.1
	0.025	0.587	0.933	0.625	0.583	-0.7
	0.0125	0.463	0.761	0.573	0.436	-5.9
	0.01	0.420	0.700	0.560	0.392	-6.7
0.75	0.05	0.983	1.327	0.749	0.994	1.2
	0.025	0.849	1.143	0.714	0.816	-3.9
	0.0125	0.677	0.932	0.683	0.636	-6.0
	0.01	0.612	0.856	0.679	0.581	-5.0
0.99	0.05	1.323	1.525	0.947	1.444	9.2
	0.025	1.132	1.313	0.952	1.251	10.5
	0.0125	0.908	1.071	0.961	1.029	13.3
	0.01	0.8248	0.984	0.964	0.948	15.0

[1] Standard errors over 100 replicates ranged from 0.003 to 0.001 for h_o^2=0.01, from 0.010 to 0.006 for h_o^2=0.25, from 0.0174 to 0.007 for h_o^2=0.50, from 0.018 to 0.009 for h_o^2=0.75 and from 0.016 to 0.009 for h_o^2=0.99. The higher and lower bound of each range corresponds to ΔF =0.05 and ΔF =0.01, respectively.

[2] obtained from simulations at selection time (t=3)

[3] predicted from equation [6] in Methods

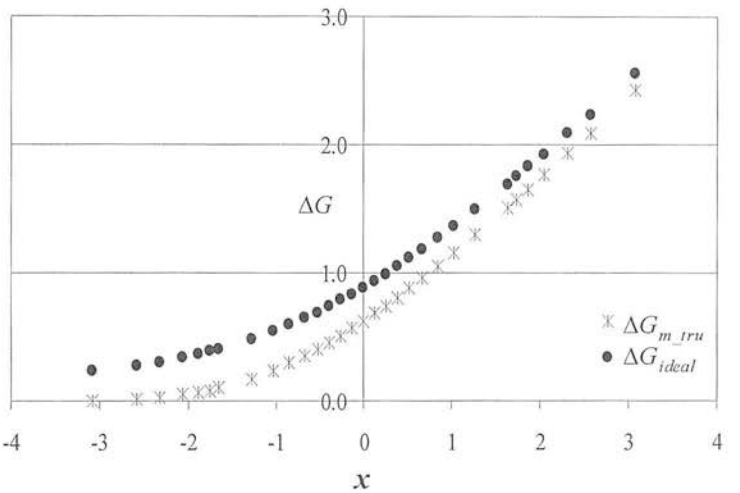
[4] predicted from equations [9] and [10] in Methods

[5] predicted from equation [8] in Methods

[6] %error =[(ΔG_{quad}- ΔG_{obs})/ ΔG_{obs}]*100

4.6. Figures

Figure 4.1. Rate of genetic gain under mass truncation selection (ΔG_{m_tru}) and theoretical ideal rate of gain (ΔG_{ideal}) after an exact allocation of long-term contributions and Mendelian sampling term for a range of values of the deviation of the truncation point from the mean (x) and of the coefficient of variance reduction (k).



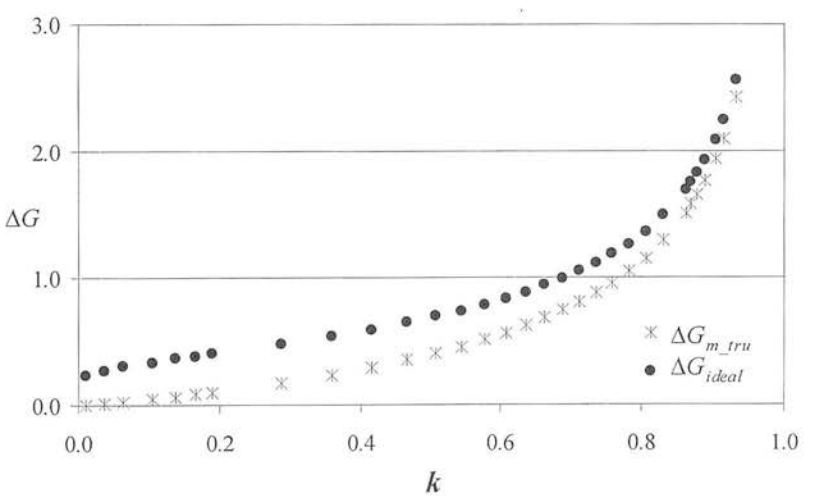


Figure 4.2. Predicted and empirical accuracy of the Mendelian sampling term (ρ) at selection and at the convergence of the long-term genetic contributions. The regression coefficients of predicted accuracy on empirical accuracy at selection and at convergence ($b_{(\rho_{sel},\rho_{sel\_emp})}, b_{(\rho_{conv},\rho_{conv\_emp})}$, respectively) and their correlation ($r_{(\rho_{sel},\rho_{sel\_emp})}, r_{(\rho_{conv},\rho_{conv\_emp})}$, respectively) are presented.

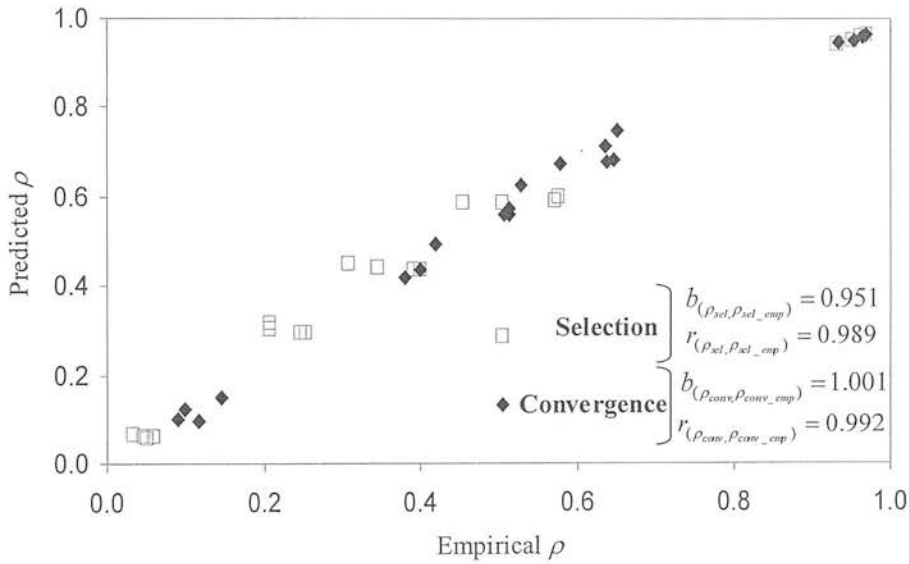
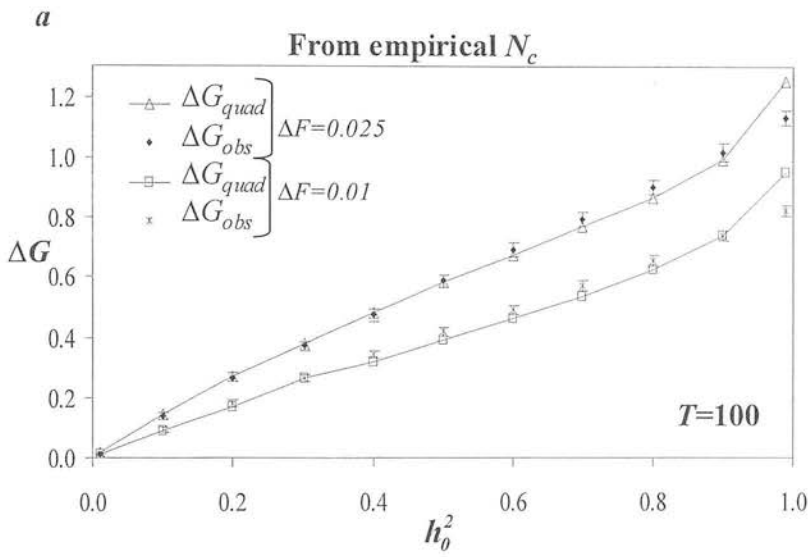


Figure 4.3. Observed (ΔG_{obs}) and predicted (ΔG_{ub}) rate of genetic gain for a range of base heritability (h_0^2) using the empirical or predicted effective number of parents at selection time (N_c), for two population sizes of T=100 and 300.



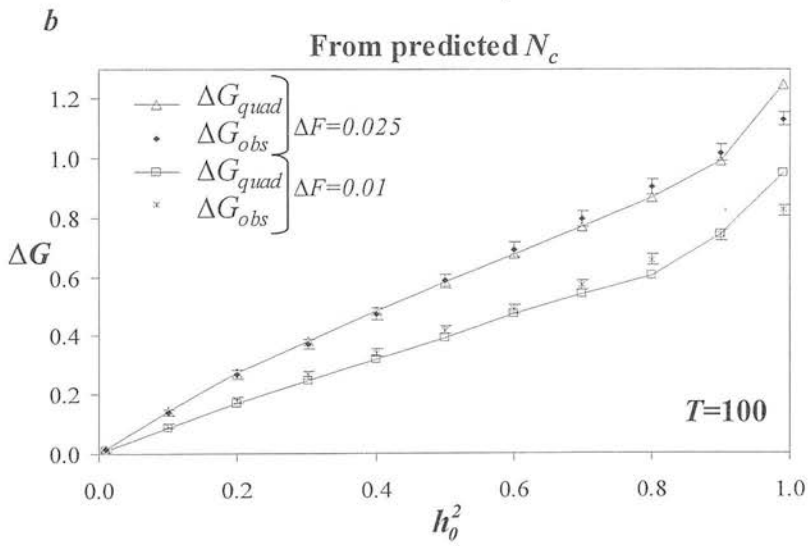
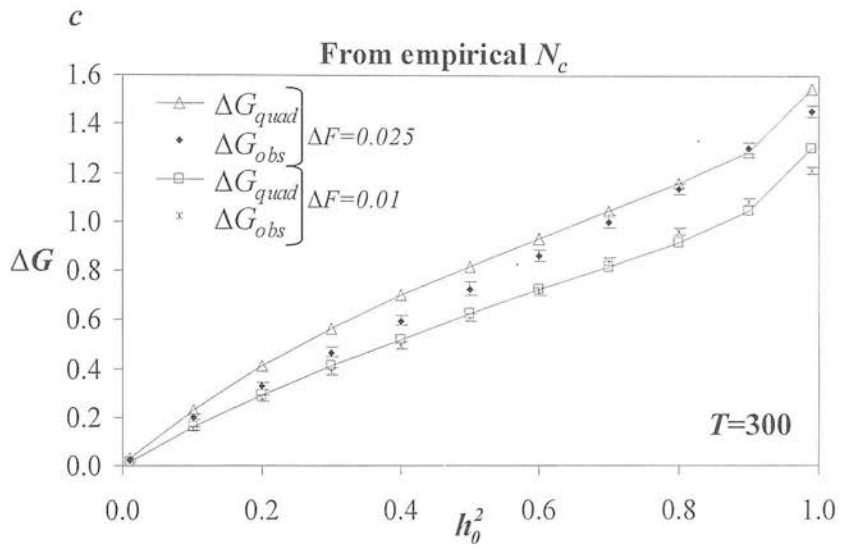


Figure 4.3 cont.



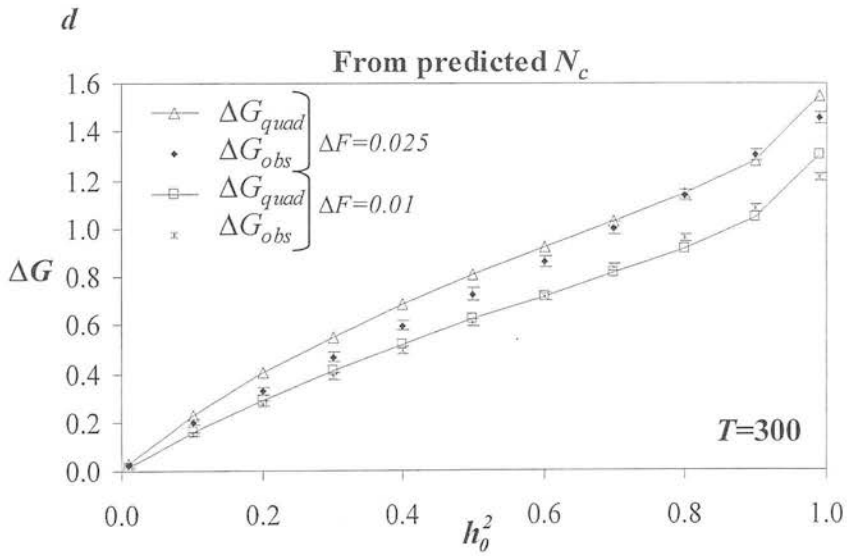


Figure 4.4. Predicted rates of gain for schemes under truncation (TRU) and quadratic optimisation (OPT) selection giving a range of ΔF (%) for two base heritabilities (h_0^2) of 0.35 and 0.10 and two population sizes of 100 and 300.

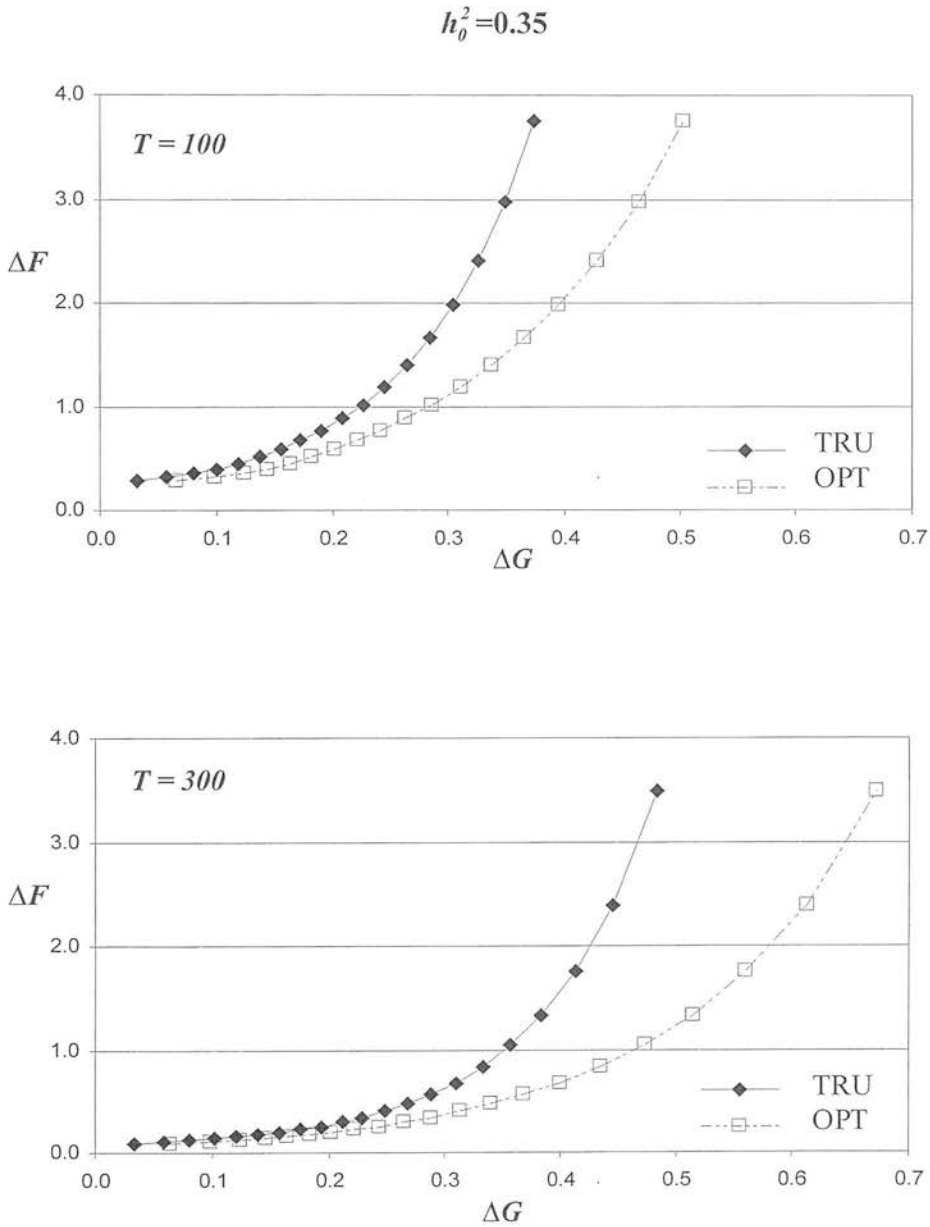
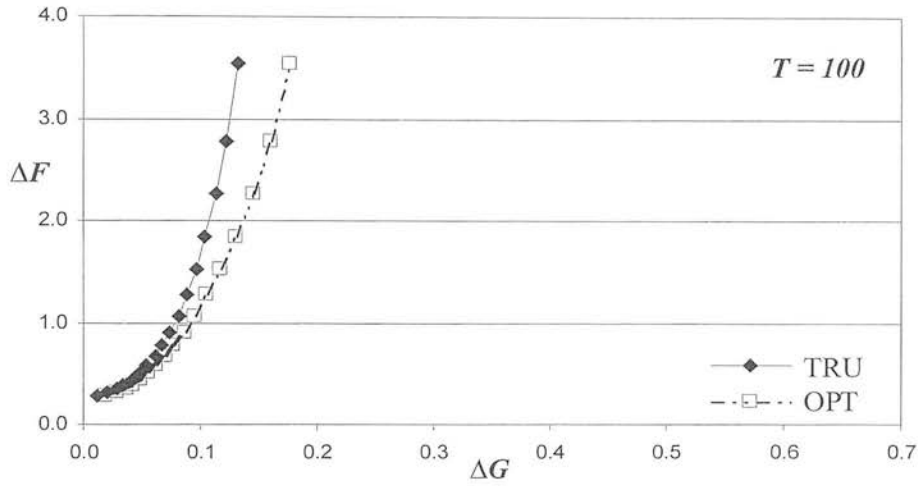


Figure 4.4 cont.

h_0^2=0.10



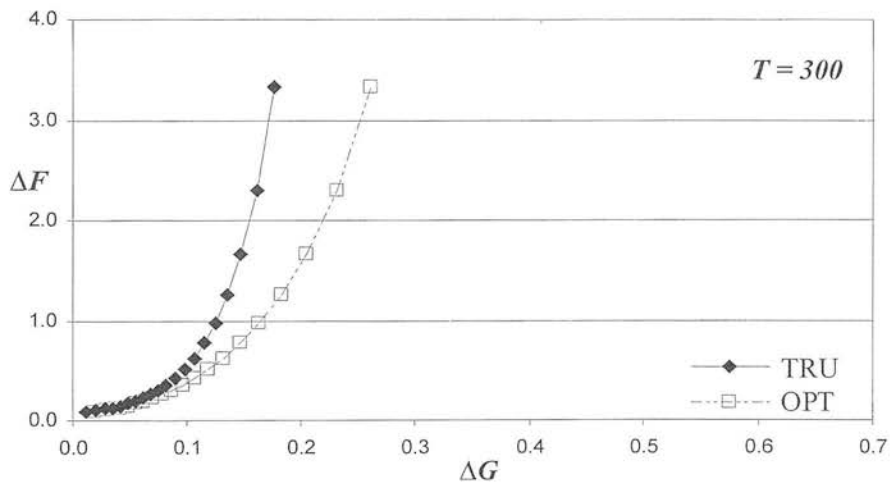
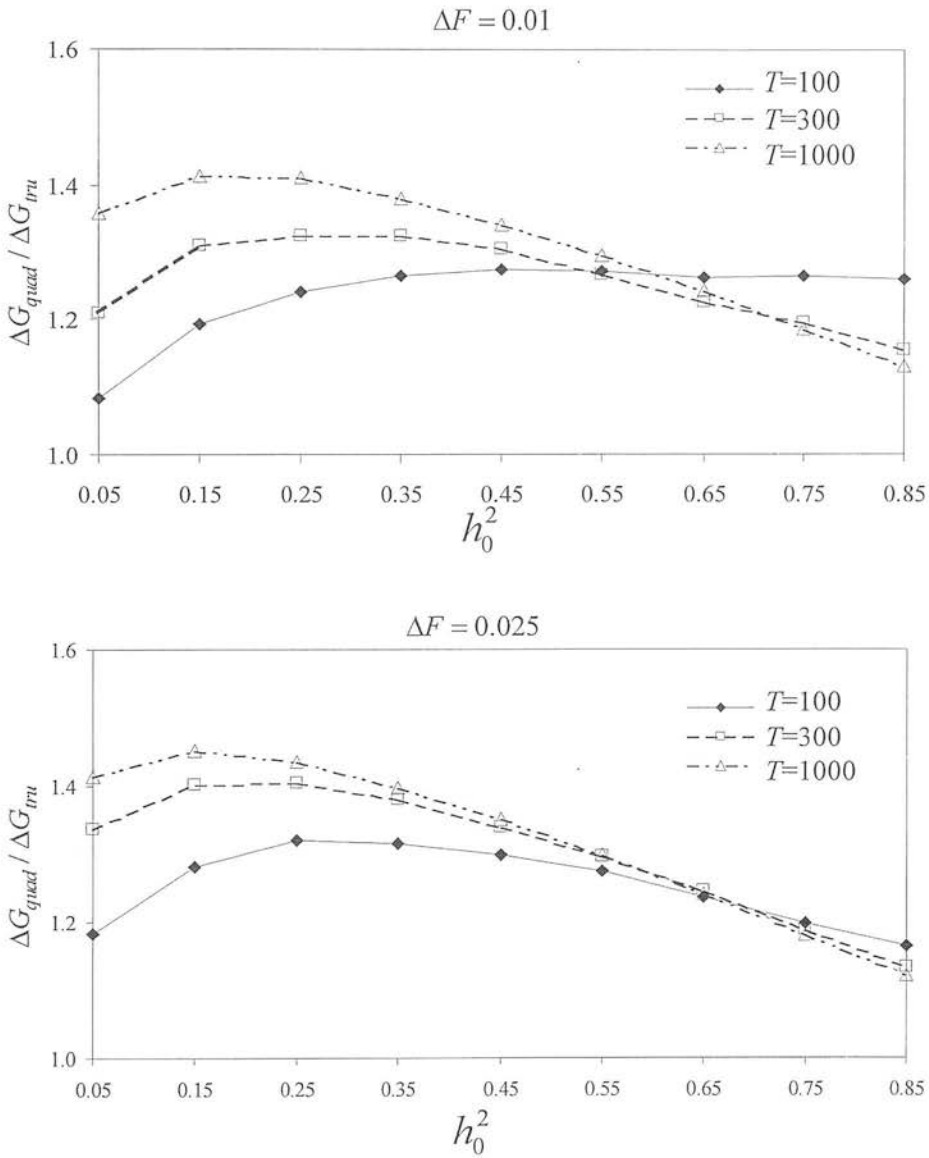


Figure 4.5. Ratio of the predicted rate of gain under quadratic optimisation (ΔG_{quad}) on the predicted gain under truncation (ΔG_{tru}) for three populations sizes (T=100, 300 and 1,000) for ΔF of 0.01 and 0.025 per generation and a range of trait base heritabilities (h_0^2).



4.7. Appendixes

Appendix 4.1. Implementation of the Newton-Raphson method for finding the truncation point and selection intensity for meeting the ΔF constraint.

The values of i and x that satisfy $(4T\Delta F)^{-1} = p(i-x)^2(1+x^2-ix)^{-1}$ are to be found.

Starting with an initial value of $x = 0.0$, iterations were performed such as

$$x_t = x_{t-1} - \frac{f(x_{t-1})}{f'(x_{t-1})}, \text{ where}$$

$$f(x_{t-1}) = \frac{p(i-x)^2}{(1-ix+x^2)} - (4T\Delta F)^{-1}, \text{ and}$$

$$f'(x_{t-1}) = \frac{2.0p(i-x)[i(i-x)-1.0]}{(1-ix+x^2)^2}$$

where p and i are taken from a Normal Distribution table from the normal deviate x_{t-1}.

For example, for T=100 and ΔF=0.01, $(4T\Delta F)^{-1} = 0.25$ The corresponding value of x that meets the restriction (i.e., the desired ΔF) is 0.2453 and the corresponding p and i are 0.403 and 0.960, respectively.

iteration	x	$f(x)$	$f'(x)$
1	0	0.0683	-0.2899
2	0.2356	0.0026	-0.2667
3	0.2453	0	-0.2656
4	0.2453	0	-0.2656

Appendix 4.2. Phenotypic and genetic variance-covariance matrix for the extended pseudo-BLUP index

- Matrix **P** of phenotypic variance-covariance among sources of information

X_i	$\overline{X}_{fs}$	$\overline{X}_{hs}$	$\hat{A}_d$	$\hat{\bar{A}}_d$	$\hat{A}_s$	$\hat{a}_d$	$\hat{\bar{a}}_d$	$\hat{a}_s$
σ^2_{Pt}	$\sigma^2_{ASt}+\sigma^2_{ADt}+\sigma^2_C+(\sigma^2_{AWt}+\sigma^2_E)/n$	$\sigma^2_{ASt}+(\sigma^2_{ADt}+\sigma^2_C)/m+(\sigma^2_{AWt}+\sigma^2_E)/mn$	$V(\hat{A}_d)/2$	$V(\hat{A}_d)/2m$	$V(\hat{A}_s)/2$	$Cov(\hat{A}_d,\hat{a}_d)/2$	$Cov(\hat{A}_d,\hat{a}_d)/2m$	$Cov(\hat{A}_s,\hat{a}_s)/2$
	$\sigma^2_{ASt}+\sigma^2_{ADt}+\sigma^2_C+(\sigma^2_{AWt}+\sigma^2_E)/n$	$\sigma^2_{ASt}+(\sigma^2_{ADt}+\sigma^2_C)/m+(\sigma^2_{AWt}+\sigma^2_E)/mn$	$V(\hat{A}_d)/2$	$V(\hat{A}_d)/2m$	$V(\hat{A}_s)/2$	$Cov(\hat{A}_d,\hat{a}_d)/2$	$Cov(\hat{A}_d,\hat{a}_d)/2m$	$Cov(\hat{A}_s,\hat{a}_s)/2$
		$\sigma^2_{ASt}+(\sigma^2_{ADt}+\sigma^2_C)/m+(\sigma^2_{AWt}+\sigma^2_E)/mn$	$V(\hat{A}_d)/2m$	$V(\hat{A}_d)/2m$	$V(\hat{A}_s)/2$	$Cov(\hat{A}_d,\hat{a}_d)/2m$	$Cov(\hat{A}_d,\hat{a}_d)/2m$	$Cov(\hat{A}_s,\hat{a}_s)/2$
			$V(\hat{A}_d)$	$V(\hat{A}_d)/m$	0	$Cov(\hat{A}_d,\hat{a}_d)$	$Cov(\hat{A}_d,\hat{a}_d)/m$	0
				$V(\hat{A}_d)/m$	0	$Cov(\hat{A}_d,\hat{a}_d)/m$	$Cov(\hat{A}_d,\hat{a}_d)/m$	0
					$V(\hat{A}_s)$	0	0	$Cov(\hat{A}_s,\hat{a}_s)$
						$V(\hat{a}_d)$	$V(\hat{a}_d)/m$	0
	Symmetric						$V(\hat{a}_d)/m$	0
								$V(\hat{a}_s)$

- Vector $\mathbf{g_A}$ (transposed) of genetic covariances between the Breeding Value of the candidate (A_i) the and nine sources of information

X_i	$\overline{X}_{fs}$	$\overline{X}_{hs}$	$\hat{A}_d$	$\hat{\bar{A}}_d$	$\hat{A}_s$	$\hat{a}_d$	$\hat{\bar{a}}_d$	$\hat{a}_s$
σ^2_{At}	$\sigma^2_{ASt}+\sigma^2_{ADt}+\sigma^2_{AWt}/n$	$\sigma^2_{ASt}+\sigma^2_{ADt}/m+\sigma^2_{AWt}/mn$	$V(\hat{A}_d)/2$	$V(\hat{A}_d)/2m$	$V(\hat{A}_s)/2$	$Cov(\hat{A}_d,a_d)/2$	$Cov(\hat{A}_d,a_d)/2m$	$Cov(\hat{A}_s,a_s)/2$

- Vector $\mathbf{g}_{a\_cand}$ (transposed) of genetic covariances between the Mendelian sampling term of the candidate (a_i) and the nine sources of information

X_i	$\overline{X}_{fs}$	$\overline{X}_{hs}$	$\hat{A}_d$	$\hat{\bar{A}}_d$	$\hat{A}_s$	$\hat{a}_d$	$\hat{\bar{a}}_d$	$\hat{a}_s$
σ^2_{at}	σ^2_{at}/n	σ^2_{at}/mn	0	0	0	0	0	0

- Vector $\mathbf{g}_{a\_sir}$ (transposed) of genetic covariances between the Mendelian sampling term of male selected candidates when they become Sires ($a_{\_sir}$) and the nine sources of information

X_i	$\overline{X}_{fs}$	$\overline{X}_{hs}$	$\hat{A}_d$	$\hat{\bar{A}}_d$	$\hat{A}_s$	$\hat{a}_d$	$\hat{\bar{a}}_d$	$\hat{a}_s$
$Cov(A_s,a_s)/2$	$Cov(A_s,a_s)/2$	$Cov(A_s,a_s)/2$	0	0	$Cov(\hat{A}_s,a_s)$	0	0	$V(\hat{a}_s)$

- Vector $\mathbf{g}_{a\_dam}$ (transposed) of genetic covariances between the Mendelian sampling term of female selected candidates when they become Dams ($a_{\_dam}$) and the nine sources of information

X_i	$\overline{X}_{fs}$	$\overline{X}_{hs}$	$\hat{A}_d$	$\hat{\overline{A}}_d$	$\hat{A}_s$	$\hat{a}_d$	$\hat{\overline{a}}_d$	$\hat{a}_s$
$Cov(A_d, a_d)/2$	$Cov(A_d, a_d)/2$	$Cov(A_d, a_d)/2m$	$Cov(\hat{A}_d, a_d)$	$Cov(\hat{A}_d, a_d)/m$	0	$V(\hat{a}_d)$	$V(\hat{a}_d)/m$	0

Appendix 4.3. Expressions used for the variances and co-variances in **P** matrix and **g** vectors after extending the pseudo-BLUP index

- **P-matrix**

At t =1 Mass selection was assumed

$Cov(\hat{A}_d, \hat{a}_d) = h^4/2$ since $Cov(\hat{A}_d, \hat{a}_d) = Cov(h^2 * p_d, 1/2h^2 * p_d)$, where p_d is the dam's phenotype.

$Cov(\hat{A}_s, \hat{a}_s) = h^4/2$

$V(\hat{a}_d) = h^4/4$

$V(\hat{a}_s) = h^4/4$

At t >1 Variances and Covariances are updated to account for selection by using k_s and k_d

$k_s=i_s(i_s-x_s)$, where i_s and x_s are the selection intensity for males and the standardized deviation of the truncation point from the male mean

$k_d=i_d(i_d-x_d)$, as for males

$$Cov(\hat{A}_s, \hat{a}_s)_t = (g_A' P^{-1} g_{a\_cand})_{t-1} - k_s \times \frac{(g_A' P^{-1} g_{a\_cand})_{t-1} \times (g_A' P^{-1} g_A)_{t-1}}{(g_A' P^{-1} g_A)_{t-1}}$$

$$Cov(\hat{A}_d, \hat{a}_d)_t = (g_A' P^{-1} g_{a\_cand})_{t-1} - k_d \times \frac{(g_A' P^{-1} g_{a\_cand})_{t-1} \times (g_A' P^{-1} g_A)_{t-1}}{(g_A' P^{-1} g_A)_{t-1}}$$

$$V(\hat{a}_d)_t = (g_{a\_cand}' P^{-1} g_{a\_cand})_{t-1} - k_d \times \frac{(g_A' P^{-1} g_{a\_cand})_{t-1}^2}{(g_A' P^{-1} g_A)_{t-1}}$$

$$V(\hat{a}_s)_t = (g_{a\_cand}' P^{-1} g_{a\_cand})_{t-1} - k_s \times \frac{(g_A' P^{-1} g_{a\_cand})_{t-1}^2}{(g_A' P^{-1} g_A)_{t-1}}$$

where $g_A' P^{-1} g_A = \sigma^2_I$

- **g-vectors**

i. $\mathbf{g}_{a\_cand}$

At t =1 Mass selection was assumed

$Cov(\hat{A}_d, a_d) = Cov(\hat{A}_d, \hat{a}_d) = h^4/2$

$Cov(\hat{A}_s, a_s) = Cov(\hat{A}_s, \hat{a}_s) = h^4/2$

At t >1 Covariances are updated to account for selection (using k_s and k_d) as shown before

ii. $\mathbf{g}_{a\_sires}$

At t =1 Mass selection was assumed

$Cov(A_s, a_s) = \sigma_a^2$, where $\sigma_a^2 = h^2/2$, since $\sigma_P^2 = 1.0$

At t>1

$$Cov(A_s, a_s)_t = \sigma_a^2 - k_s * \frac{(g_A' P^{-1} g_A)_{t-1} * (g_A' P^{-1} g_{a\_cand})_{t-1}}{(g_A' P^{-1} g_A)_{t-1}} = \sigma_a^2 - k_s \times (g_A' P^{-1} g_{a\_cand})_{t-1}$$

iii. $\mathbf{g}_{a\_dams}$

At t =1 Mass selection was assumed

$Cov(A_d, a_d) = \sigma_a^2$

At t>1

$$Cov(A_d, a_d)_t = \sigma_a^2 - k_d * \frac{(g_A' P^{-1} g_A)_{t-1} \times (g_A' P^{-1} g_{a\_cand})_{t-1}}{(g_A' P^{-1} g_A)_{t-1}} = \sigma_a^2 - k_d \times (g_A' P^{-1} g_{a\_cand})_{t-1}$$

Appendix 4.4. Prediction of the effective number of parents at the time of selection (N_c)

Figure 4.4.1. Observed relationship between the empirical ratio N_c/N_r and $1-h_0^2$, for eight levels of the joint factor $T\Delta F$ taken from T=100, 200 and 300, and ΔF=0.025, 0.0125 and 0.010.

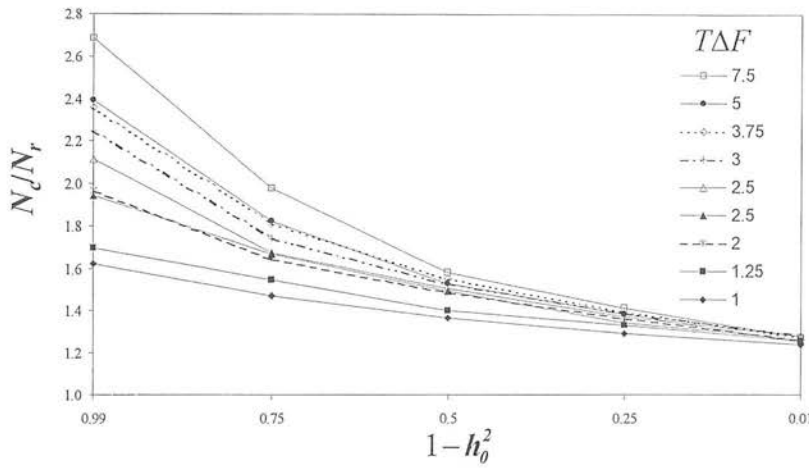


Figure 4.4.2. Relationship between the double natural logarithm of the ratio N_c/N_r ($\ln[\ln(N_c/N_r)]$ and $1-h_0^2$, for eight levels of the joint factor $T\Delta F$ taken from T=100, 200 and 300, and ΔF=0.025, 0.0125 and 0.010.

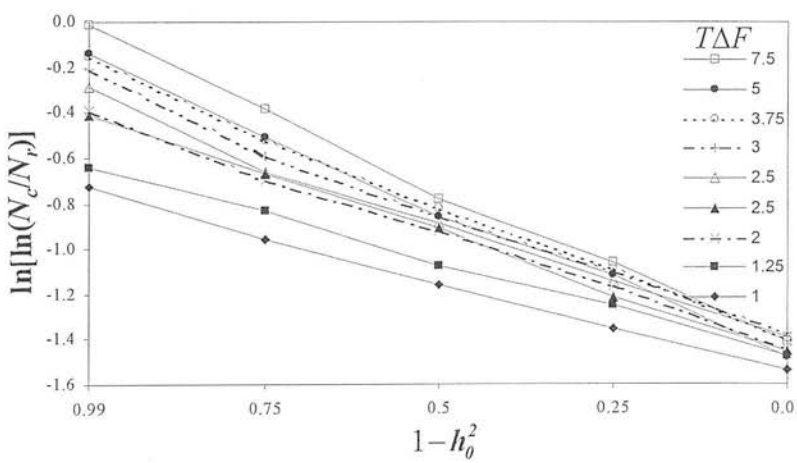
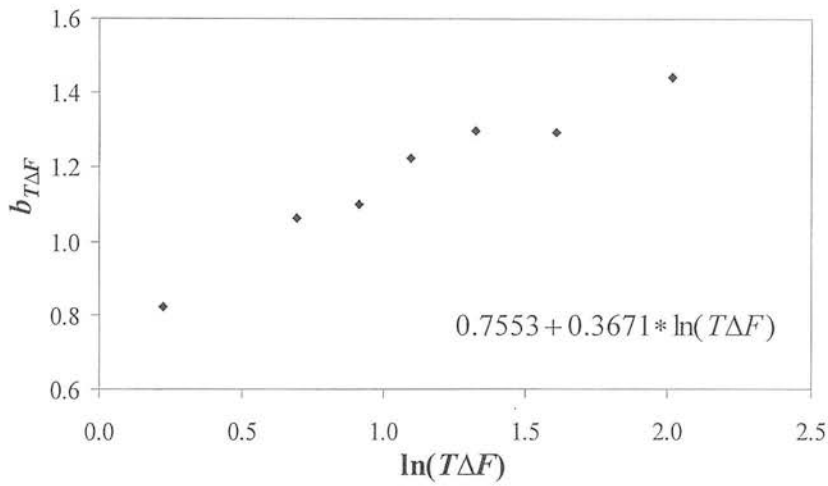


Figure 4.4.3. Regression of the regression coefficients of $\ln[\ln(N_c/N_r)]$ for each level of the joint factor $T\Delta F$ (see Figure 2) on the natural logarithm of $T\Delta F$ ($\ln(T\Delta F)$).



The $\ln[\ln(N_c/N_r)]$ can be predicted as:

$$\ln[\ln(N_c/N_r)] = -1.459 + [0.755 + 0.367 \times \ln(T\Delta F)] \times (1 - h_0^2)$$

where the intercept corresponds to the common intercept of the nested regression of $\ln[\ln(N_c/N_r)]$ on $1 - h_0^2$ for each level of $T\Delta F$.

Taking anti-logs the above expression can be simplified to:

$$\ln(N_c/N_r) = 0.2325 \times (T\Delta F)^{0.3671} \times e^{[0.7553(1-h_0^2)]}$$

5. CHAPTER FIVE

Potential Benefits of using Identified Genes and Quadratic Indices in Two-Trait Breeding Goals

5.1. Introduction

Breeding goals in selection programmes for livestock populations commonly include several traits. However, the vast majority of the research directed to evaluate the potential benefits of using information on identified quantitative trait loci or QTL (GAS) or on markers linked to them (MAS) in selection decisions has been focused on single-trait scenarios (e.g., for GAS: Villanueva *et al.*, 1999; Abdel-Azim and Freeman, 2002; for MAS: Ruane and Colleau, 1995 and 1996; Meuwissen and Goddard, 1996; Villanueva *et al.*, 2002b).

The benefits of GAS or MAS over phenotypic selection (i.e., selection ignoring molecular information) in multi-trait scenarios have been evaluated only in a small number of studies. de Koning and Weller (1994) studied the benefits from GAS when selecting on an index that included two traits affected by a codominant biallelic QTL. The highest relative selection efficiency of GAS over phenotypic selection was found when the traits were negatively correlated. As in single trait scenarios, the benefits from the use of molecular information increased as the proportion of the genetic variance explained by the QTL increased and the heritability of both traits decreased. The gain from GAS was up to 4.5 times the gain from phenotypic selection. Xie and Xu (1998) used the approach of Lande and Thompson (1990) to deterministically predict the relative efficiency of MAS over phenotypic selection for the improvement of two or four positively correlated traits. The gain from MAS was up to 2.8 (with two traits) or 5.0 (with four traits) times the gain from conventional index selection ignoring marker information. Whilst these studies suggest potential substantial benefits from using molecular information in multiple breeding goals, they assumed mass selection (i.e., the estimated breeding values were obtained from own phenotype alone) rather than best linear unbiased prediction (BLUP) of breeding values.

Verrier (2001) used the MAS-BLUP methodology of Fernando and Grossman (1989) and investigated the value of MAS in outbred populations when selection is on an index that included two negatively correlated traits. He found extra short-term gains

in the aggregate genotype from MAS (i.e., more than 5.0% relative to conventional BLUP selection up to generation five of selection) when the QTL had positive pleiotropic effects on both traits, both traits had equal relative economic weight, and males had no records for one of the traits. This extra short-term gain from MAS arose from an increased gain in the trait with missing data, but at the expense of overall accumulated response for the aggregate genotype in the long-term (i.e., from generation 5 onwards, the gain from MAS was 1.0 to 2.5% lower than the gain from conventional selection). No extra gains in the aggregate genotype were found from MAS when the QTL affected only the trait with lower relative weight and phenotypes for both traits were available for both sexes. In this case, the only benefit from MAS over conventional selection arose from avoiding, to a greater extent, the loss of the favourable allele, but only when the QTL had a large effect and the marker was close to the QTL.

All the studies mentioned above have assumed standard truncation selection with fixed contributions of selected candidates. This could hamper the comparison between selection schemes using and ignoring QTL information as they may lead to different inbreeding levels.

Under a single trait scenario with GAS, Villanueva *et al.* (1999) used a dynamic selection tool which optimises genetic contributions of selection candidates for obtaining maximum genetic gain while constraining the rate of inbreeding to a pre-defined value. The dynamic optimisation tool is implemented as a quadratic index in which the desired inbreeding rate is achieved by applying a quadratic constraint on the average co-ancestry of selection candidates weighted by their projected use (e.g., Woolliams *et al.*, 2002). Quadratic optimisation selection in GAS schemes allowed for increased genetic gains when compared to truncation selection at the same rate of inbreeding. Villanueva *et al.* (2002b) extended the optimisation methods to MAS but also under a single trait scenario.

The objective of this study was to investigate, through stochastic simulation, the benefits from GAS when the optimisation method with constrained inbreeding is

applied on an index including two correlated traits, only one of which is affected by an identified QTL. This would help to address the relevant issue in practical breeding schemes of how to include identified QTL in current selection decisions, and will describe the impact on the overall breeding goal and on individual traits. For instance, a realistic example would be the simultaneous selection for blood oxygen saturation levels (indicator trait for susceptibility to ascites) and growth related traits in broilers, with a QTL for the former (Navarro *et al.*, 2002). This study focused on a scenario in which benefits from the use of QTL information might be anticipated: the traits were negatively correlated and the trait affected by the QTL had a low heritability. Situations where both sexes had phenotypes available for both traits, or where the phenotypic expression of the trait affected by the QTL was sex-limited were evaluated.

5.2. Methods

Selection schemes using or ignoring information on the QTL when estimating the total breeding value were compared using stochastic simulations. Two selection procedures were used: i) standard truncation selection (T) in which a fixed number of candidates from each sex (those with the highest index values) are selected each generation; and ii) optimal selection (O) in which the number of selected candidates and their contributions are optimised to maximise genetic gain while restricting the rate of inbreeding (ΔF) per generation (Grundy *et al.*, 1998a; Villanueva *et al.*, 1999). A total of 500 replicates were run for each simulation.

5.2.1. Genetic model

One of the traits (t_1) was controlled only by polygenes (i.e., the infinitesimal model) while the other trait (t_2) was controlled by an identified additive biallelic QTL (alleles A and B) and by polygenes (i.e., a mixed inheritance model). The total genetic values of the i^{th} individual for t_1 and t_2 were respectively $g_{1_i} = u_{1_i}$ and

$g_{2_i} = u_{2_i} + v_i$, where u_{k_i} is the polygenic effect for trait k (k = 1, 2) and v_i is the genotypic value due to the QTL. The polygenic plus environmental variances summed to one for both traits. The polygenic heritabilities (h^2) were 0.3 and 0.1 for t_1 and t_2, respectively. The genotypic value due to the QTL was a, 0 and $-a$ for genotypes AA, AB and BB, respectively so a is defined as half the difference between the two homozygotes (Falconer and Mackay, 1996). For t_2, the additive genetic variance explained by the QTL (σ_v^2) was $2p(1\text{-}p)a^2$, where p is the frequency of the favourable allele (Falconer and Mackay, 1996). The initial p was 0.15. The proportion of the total genetic variation explained by the QTL in t_2 (i.e., $\theta = \sigma_v^2 / (\sigma_v^2 + \sigma_{u_2}^2)$) was 0.1 or 0.5, which correspond to a values (i.e., $a = \sqrt{\theta\sigma_{u_2}^2 / [(1-\theta)2p(1-p)]}$) of 0.21 and 0.63 units of phenotypic standard deviation (σ_P). The polygenic genetic correlation (ρ_u) between both traits was -0.5 and no environmental correlation was assumed.

5.2.2. Simulation of the population

The base generation (t = 0) was composed of N = 120 individuals (60 males and 60 females) with family structure. A prior randomly mated generation (t = -1) was simulated to create the family structure in t = 0. Generation 1 (t = 1) was obtained from the matings among selected individuals at t = 0. In the base generation the polygenic values (u_1 and u_2) were drawn from a bivariate normal distribution with correlation ρ_u and polygenic variances $\sigma_{u_1}^2$ and $\sigma_{u_2}^2$. Polygenic genetic values in t = 0 for the i^{th} individual were generated as $u_{1_i} = \sigma_{u_1} w_1$ and $u_{2_i} = [\rho_u w_1 + \sqrt{(1-\rho_u^2)} w_2]\sigma_{u_2}$, where w_1 and w_2 are random normal deviates taken from a normal distribution with variance equal to one. Phenotypic values for t_1 (y_{1_i}) and t_2 (y_{2_i}) were obtained by adding an environmental component to the total genetic value. Selection was carried out for 10 discrete generations. From t = 1 to 10, the polygenic value for trait k (k = 1, 2) was generated as the parental average polygenic

value plus a random Mendelian sampling term. The latter was sampled from a bivariate normal distribution (i.e., in the same way as for the polygenic values in t = 0) with correlation ρ_u and variances $\frac{1}{2}\sigma^2_{u_k}(1-F)$ for $k = 1,2$, where F is the average parental inbreeding coefficient. The genotype for the QTL was obtained by randomly sampling one allele from each parent. The number of selection candidates was kept constant across generations.

5.2.3. Estimation of breeding values

5.2.3.1. Schemes using QTL information

In schemes using the QTL genotype (denoted as G) it was assumed that the QTL effect and QTL genotypes for all individuals were known without error. The total estimated breeding values for t_1 and t_2 were respectively $EBV_{t_1} = EBV_{u_1}$ and $EBV_{t_2} = EBV_{u_2} + BV_{qtl}$, where EBV_{u_k} is the estimate of the polygenic breeding value for trait k and BV_{qtl} is the known breeding value due to the QTL effect. The BV_{qtl} was $2(1\text{-}p)a$, $[(1\text{-}p)\text{-}p]a$ and $-2pa$ for genotypes AA, AB, BB, respectively (Falconer and Mackay, 1996), with p updated each generation.

A bivariate BLUP evaluation was performed to obtain EBV_{u_1} and EBV_{u_2} simultaneously using PEST (Groeneveld *et al.*, 1990). In G schemes, EBV_{u_k} was obtained using the base population polygenic variances ($\sigma^2_{u_1}$ and $\sigma^2_{u_2}$), and the phenotypic values for t_2 corrected for the QTL effect ($y^*_{2_i} = y_{2_i} - v_i$).

5.2.3.2. Schemes ignoring QTL information

In those schemes ignoring the QTL genotypes (denoted as I), EBV_{t_1} and EBV_{t_2} were obtained from a bivariate BLUP genetic evaluation performed using the base

population total genetic additive variances ($\sigma^2_{g_1} = \sigma^2_{u_1}$ and $\sigma^2_{g_2} = \sigma^2_{u_2} + \sigma^2_v$) and the phenotypic values y_{2_i} uncorrected for the QTL effect.

5.2.4. Availability of phenotypes in selection candidates

Two situations were simulated with regard to phenotypes availability: i) all selection candidates had available phenotypes for both traits (termed as *full-data*), and ii) all selection candidates had available phenotypes for t_1, but only female candidates had available phenotypes for t_2 (termed as *sex-limited*).

5.2.5. Selection schemes

The breeding goal was $H = BV_1 + BV_2$, where BV_k is the true breeding value for trait k, which includes only the polygenic component (t_1), or the polygenic plus the QTL component (t_2) according to the genetic models described before. Thus, both traits had the same economic weight. The index on which selection was carried out was $I_S = EBV_{t_1} + EBV_{t_2}$.

Four schemes were evaluated depending on the selection method and the use of QTL information: truncation ignoring (T_I) or using (T_G) the QTL genotypes and optimal selection ignoring (O_I) or using (O_G) the QTL genotypes. Random mating among selected candidates was performed for both truncation and optimal selection.

5.2.5.1. Truncation selection

In truncation selection schemes, the 12 highest ranked male and females based on the index I_S were selected each generation and each selected individual was mated only once. All females produced the same number of offspring (i.e., ten), hence equal contributions to the next generation were allocated to each selected candidate.

5.2.5.2. Optimal selection

In optimal schemes the number of selected male and female candidates and their contributions to the next generation were optimised for maximising the rate of gain (ΔG) while the rate of inbreeding was constrained to a pre-defined level each generation (Meuwissen, 1997; Grundy *et al.*, 1998a). The optimisation algorithm described by Meuwissen (1997) was used. An explicit optimum was found by maximising the following objective function:

$$H_{quad} = \mathbf{c}_t^T \mathbf{g}_t - \lambda_0(\mathbf{c}_t^T \mathbf{A}_t \mathbf{c}_t - C_t) - [\mathbf{c}_t^T \mathbf{Q} - (1/2)\mathbf{1}^T]\boldsymbol{\lambda}$$

where $\mathbf{c}_t$ is the vector of contributions of selected candidates at generation t to the next generation, $\mathbf{g}_t$ is the vector of BLUP-EBV, $\mathbf{A}_t$ is the additive relationship matrix among selection candidates, $C_t = 2[1-(1-\Delta F)^t]$ (Grundy *et al.*, 1998a), $\mathbf{Q}$ is a known incidence matrix for sex, $\mathbf{1}$ equals [1 1] , and λ_0 and $\boldsymbol{\lambda}$ are Lagrangian multipliers. The third term in H_{quad} ensures that male and female parents contribute with a half of the gene pool each. The output of the optimisation procedure is the vector of optimum mating proportions $\mathbf{c}$ (i.e., contributions) of candidates at any particular generation. The optimal number of offspring for an individual i is $2Nc_i$ (a real number), and the actual (integer) number of offspring per parent was obtained following Grundy *et al.* (1998a). Selected candidates are those with $c_i > 0$ and they contribute to the next generation according to their c_i values. Since the pre-defined ΔF is achieved by a setting a constraint on a quadratic form on the candidates projected usage (i.e., $\mathbf{c}_t^T \mathbf{A}_t \mathbf{c}_t \leq C_t$), this dynamic optimisation tool receives the generic denomination of quadratic index (Woolliams *et al.*, 2002).

With optimal selection, ΔF was restricted to the value obtained in the corresponding truncation schemes. It must be noted that here, the input for the optimisation algorithm are the I_s values of selection candidates, rather than the EBV of individual traits as in previous studies under single trait scenarios with polygenic (Meuwissen,

1997; Grundy *et al.*, 1998a), or mixed inheritance models (Villanueva *et al.*, 1999 and 2002b).

5.2.6. Bias in BLUP evaluation when QTL information is ignored

There is bias in the estimation of breeding values when the QTL information is ignored and phenotypic records uncorrected for the QTL effect are used. In addition, specifically with BLUP evaluation there is an extra bias when the QTL is ignored due to the fact that the EBV of an individual is regressed (incorrectly) toward its parents performance. This leads to different bias among candidates with the same genotype (as it depends on the genotype of the parents) which causes additional ranking errors within genotypes (Villanueva *et al.*, 1999). The bias in the BLUP evaluation for trait k was calculated as $EBV_k - u_k$ for the four schemes evaluated.

5.2.7. Gamete phase disequilibrium between the QTL and polygenes

Selection creates gamete phase disequilibrium by promoting that gametes carrying alternative QTL alleles have different polygenic values (e.g., Kennedy *et al.*, 1992; Dekkers and van Arendonk, 1998). The disequilibrium has been previously identified as a leading factor affecting the extent of the so called short term *versus* long-term conflict when selection is applied on traits under mixed inheritance (e.g., Gibson, 1994; Pong-Wong and Woolliams, 1998; Dekkers and van Arendonk, 1998; Villanueva *et al.*, 1999).

The disequilibrium between the QTL alleles and the polygenes was modelled at the gametic level following the approach of Dekkers and van Arendonk (1988). The disequilibrium for trait k was computed as $(\overline{u}_{B_k} - \overline{u}_{A_k})/(\sigma^2_{g_k} + \sigma^2_{e_k})$, where $\overline{u}_{B_k}$ and $\overline{u}_{A_k}$ are the average polygenic values of gametes which form generation t containing the unfavourable and the favourable QTL allele, respectively. Let f_{AA_k}, f_{AB_k} and f_{BA_k} be respectively the frequencies of selected parents with genotypes AA, AB and

BA among those parents carrying allele A (i.e., $f_{AA_k} + f_{AB_k} + f_{BA_k} = 1$) and let $\bar{u}_{AA_k}$, $\bar{u}_{AB_k}$ and $\bar{u}_{BA_k}$ be the corresponding average polygenic values for the three genotypes. Parents AA will produce 100% of gametes A with breeding value $\bar{u}_{AA_k}/2$, parents AB will produce 50% of gametes A with breeding value $\bar{u}_{AB_k}/2$ and parents BA will produce 50% of gametes A with breeding value $\bar{u}_{BA_k}/2$. Hence, the average polygenic value of gametes carrying the favourable allele A is

$$\bar{u}_{A_k} = [2f_{AA_k}(\tfrac{1}{2}\bar{u}_{AA_k}) + f_{AB_k}(\tfrac{1}{2}\bar{u}_{AB_k}) + f_{BA_k}(\tfrac{1}{2}\bar{u}_{BA_k})]/(2f_{AA_k} + f_{AB_k} + f_{BA_k}).$$

Likewise, for gametes carrying the allele B,

$$\bar{u}_{B_k} = [2f_{BB_k}(\tfrac{1}{2}\bar{u}_{BB_k}) + f_{AB_k}(\tfrac{1}{2}\bar{u}_{AB_k}) + f_{BA_k}(\tfrac{1}{2}\bar{u}_{BA_k})]/(2f_{BB_k} + f_{AB_k} + f_{BA_k}),$$ where now frequencies and average polygenic values are computed among the selected parents carrying allele B (i.e., $f_{BB_k} + f_{AB_k} + f_{BA_k} = 1$).

5.3. Results

5.3.1. *Fixation and loss of the favourable allele*

The gene frequency of the favourable allele and the percentage of replicates in which it was lost for scenarios with phenotypic records in both sexes and both traits or with sex-limited phenotypes for t_2 are presented in Figure 5.1. As expected, allele fixation was faster for $\theta = 0.5$. Schemes using the gene information (O_G and T_G) fixed ($p \geq 0.97$) the favourable allele within the selection period considered and gave similar rates of increase in p in both the *full-data* and *sex-limited* scenarios. Fixation was faster in O_G (at $t = 4$ and $t = 2$ for $\theta = 0.1$ and $\theta = 0.5$, respectively) than in T_G (at $t = 7$ and $t = 3$ for $\theta = 0.1$ and $\theta = 0.5$, respectively). In contrast, schemes ignoring the gene information (O_I and T_I) did not lead to fixation of the favourable allele within the selection period evaluated (i.e., ten generations), although p was close to fixation in the *full-data* scenario for $\theta = 0.5$. Although not always significant, a trend was identified in which O_I increased p faster than T_I during initial generations, whereas

the opposite was true for later generations. This was particularly evident for the *full-data* scenario for $\theta = 0.1$ and for both scenarios for $\theta = 0.5$.

The loss of the favourable allele was only important in those schemes ignoring the gene information, particularly for $\theta = 0.1$, and for the *sex-limited* scenario (Figure 5.1). Also, the O_I scheme consistently lost the favourable allele to a greater extent than T_I.

5.3.2. Response in the breeding objective (H)

5.3.2.1. Phenotypic records available in both sexes for both traits

Table 5.1 shows the accumulated response in H for the four schemes evaluated under the *full-data* scenario. The ΔF in optimal schemes (O_I and O_G) was restricted to 0.06 per generation, which was the value observed in truncation schemes.

For a particular selection scheme, the gain in the breeding objective was always greater for $\theta = 0.5$ than for $\theta = 0.1$ at all generations of selection. At each generation, schemes using the genotype information (O_G and T_G) yielded more gain than the corresponding schemes ignoring the QTL (O_I and T_I) at both levels of θ. The advantage of G schemes was highest before fixation (around 12% for $\theta = 0.1$ and 50% for $\theta = 0.5$ at t=2) and decreased after fixation (around 6% for $\theta = 0.1$ and 2% for $\theta = 0.5$ by t=10).

Optimal selection schemes (O_G and O_I) achieved higher gain in H than the corresponding truncation selection schemes (T_G and T_I) at both levels of θ during the whole selection period. By $t = 10$, the accumulated gain in H was about 12% higher for O_G than for T_G and about 13% higher for O_I than for T_I for both θ levels.

The comparison between O_I and T_G schemes indicates that the optimisation of contributions had a higher impact on genetic gain than the use of QTL information. The scheme O_I achieved higher accumulated gain in H than T_G during the whole selection period for $\theta = 0.1$. For $\theta = 0.5$ T_G achieved higher gains than O_I until $t = 3$, when the favourable allele is fixed in T_G. After fixation (i.e., from $t = 4$ onwards) O_I progressively overperformed T_G achieving 10% higher gain in H by generation ten. Therefore, in early generations while the gene is segregating in the T_G scheme, the importance of the use of molecular information relative to the optimisation of contributions was dependent on the size of the gene. To help disentangling the causes of this observation, a further analysis of the consequences of the use of the gene in the individual traits included in H will be presented in a later section.

The combination of optimal contributions and genotype information (i.e., the O_G scheme) resulted in the highest gains in H in both the short and the long-term at both levels of θ. When compared to traditional truncation selection (T_I), the O_G scheme gave 19% and 14% higher cumulated gain in H at $t = 10$ for $\theta = 0.1$ and $\theta = 0.5$, respectively.

5.3.2.2. Phenotypic records for trait 2 only available on females

When only females had available phenotypes in the trait controlled by the QTL (t_2) the same ranking among selection schemes to that described when both sexes had available phenotypes in both traits was observed (Table 5.2). Although the absolute values for the cumulated gains in H were lower for the *sex-limited* scenario than for the *full-data* scenario (see also Table 5.1) the differences were relatively small. At $t = 10$, the decrease in accumulated gain in H for the *sex-limited* scenario relative to the *full-data* scenario across the four schemes ranged from 0.6% to 2.7% for $\theta = 0.1$ and from 0.6% to 5.6% for $\theta = 0.5$.

The optimisation of contributions had similar benefits to those found in the *full-data* scenario (i.e., the same advantage of around 11% of O_G over T_G and of around 13% of O_I over T_I at t=10 for both values of θ were observed in both scenarios).

The advantage of *G* schemes over *I* schemes in accumulated gain in *H* at t=10 was 8.4% (T_G over T_I) and 4.3% (O_G over O_I) for θ= 0.1 and 8.0% (T_G over T_I) and 6.1% (O_G over O_I), respectively. This indicates a greater benefit of using the QTL genotypes information in the *sex-limited* scenario compared to the *full-data* scenario except for optimised selection with θ = 0.1. As in the *full-data* scenario, the O_G scheme achieved the highest cumulated gains in *H* in the short and long-term for both levels of θ. For the *sex-limited* scenario, the advantage of O_G over truncation selection ignoring the QTL (T_I) at t=10 was 20% for both θ levels. This represents a 6% increase for the same comparison under the *full-data* scenario for θ= 0.5.

5.3.3. Total and polygenic responses in individual traits

The genetic gains for the individual traits included in the breeding objective are presented below for each of the four schemes studied. Comparisons between schemes are presented only for the full-data scenario as similar results were found for the sex-limited trait scenario.

5.3.3.1. Genetic gains in trait 1

Figure 5.2 shows the total gain in t_1 (i.e., polygenic gain) across generations. Optimal schemes (O_I and O_G) clearly yielded higher gains than truncation schemes (T_I and T_G) throughout the whole selection period for both levels of θ.

For θ = 0.1, schemes ignoring QTL information achieved higher gain than the corresponding schemes using this information, although differences between schemes were small (e.g., about 2% at t=10). In contrast, for θ = 0.5 the ranking

between G and I schemes depended upon the frequency of the favourable allele. In early generations while the allele is segregating in all the schemes, the O_I and T_I schemes achieved higher gains in t_1 than the corresponding O_G and T_G schemes. Conversely, after fixation in the G schemes (i.e., from $t = 4$), these schemes achieved higher accumulated polygenic gain in t_1 than their corresponding I schemes and the extra gains ranged from 11% to 13% at t=10. Thus for $\theta = 0.5$, the lower (polygenic) gain of the O_G and T_G schemes while the gene is segregating suggests negative consequences in t_1 from selecting on the QTL affecting t_2 (even though t_1 had a complete polygenic inheritance), but not to a degree to which the long-term gain in t_1 would be compromised.

5.3.3.2. Genetic gains in trait 2

Figure 5.3 shows the total gain (i.e., polygenic plus QTL gain) and the polygenic gain for t_2. When the QTL had a small effect (i.e., $\theta = 0.1$), the total response in t_2 in schemes using QTL information was only positive after t=1 and while the favourable allele was segregating (i.e., up to t=4 in O_G and up to t=6 in T_G). After fixation the total gain in t_2 was negative. Schemes ignoring the QTL led to negative total response in t_2 throughout the whole selection period. The negative response was clearly driven by the negative sign of ρ_u. Hence, the T_G scheme was the best with respect to t_2 as it allowed positive total responses for a longer period and gave lower negative responses after fixation when compared to O_G and the schemes ignoring the QTL. As for t_1, with $\theta = 0.1$ the O_I scheme achieved the highest total response in t_2, although in the undesired direction.

With $\theta = 0.1$, the polygenic gain in t_2 followed the same trend that of t_1 (see also Figure 5.2) but in opposite direction (i.e., negative). Optimal schemes (O_G and O_I) achieved higher gains than truncation selection schemes (T_G and T_I) and the differences between schemes using or ignoring the QTL information were relatively small. Therefore, it should be emphasized that for a gene of small effect, the trends in

polygenic responses in both traits were practically determined only by the selection method, that is, truncation or optimal selection.

When the QTL had a large effect (i.e., $\theta = 0.5$) the four schemes evaluated yielded positive total responses in t_2 during the whole selection period (Figure 5.3). The highest responses were observed while the gene was still segregating. The maximum accumulated gain in O_G and T_G was observed at generations two and three, respectively, which are the generations where the QTL was fixed (see also Figure 5.1). After fixation the total response decreased steadily in both schemes although at a higher rate in the O_G scheme. Schemes ignoring the QTL showed increasing cumulative gains during the whole selection period as the favourable allele was still segregating in the last generation of selection (see also Figure 5.1).

As for the case where the QTL had a smaller effect (i.e., $\theta = 0.1$), negative polygenic gains were observed in t_2 for all selection schemes and these were higher with optimal than with truncation selection. However, in contrast with the results for $\theta = 0.1$, for $\theta = 0.5$ *G* schemes (O_G and T_G) achieved significant higher polygenic negative responses than their corresponding *I* schemes (O_I and T_I) both in the short and the long-term. Thus, in this case, both the selection method and the use of genotype information determined the ranking among schemes relative to polygenic gain.

Summarising, for $\theta = 0.1$, optimal schemes achieved the greatest accumulated polygenic responses in both traits, positive in t_1 and negative in t_2 (see both Figures 5.2 and 5.3) and the differences between schemes using or ignoring the QTL were very small. For $\theta = 0.5$, the combination of both optimal selection and the use of the QTL information (i.e., the O_G scheme) yielded the greatest accumulated polygenic response in both traits, again positive in t_1 and negative in t_2.

The reason for the T_G scheme achieving higher gains in *H* than O_I before fixation for $\theta = 0.5$ (i.e., $t = 2$ and 3, see Table 5.1) is explained by the notable increase in total gain in t_2 clearly driven by QTL gain (see Figure 5.3). As mentioned above, after

fixation in T_G the total gain in t_2 was reduced, while in O_I the gene segregates throughout the whole selection period, thus, allowing an increase in total gains in t_2 up until $t = 10$. In addition, the polygenic gains in t_1 for O_I were always higher than for T_G (see Figure 5.2) hence, allowing higher gains in H after fixation.

5.3.4. Bias in BLUP evaluation

The bias in the BLUP evaluation for the four selection schemes evaluated and two values of θ is shown for the full-data scenario in Figure 5.4. For $\theta = 0.1$ there was no significant bias in t_1 in none of the four schemes evaluated. In contrast, both T_I and O_I showed increasing levels of upward bias in t_2 across generations. By $t = 10$, T_I led to a greater bias than O_I (0.087 and 0.054 σ_p units, respectively). For $\theta = 0.5$, there was a significant downward bias in t_1 in both T_I and O_I schemes throughout the ten generations of selection (i.e., around 0.025 σ_p units for both schemes). Striking levels of positive bias in t_2 were observed for both T_I and O_I schemes, reaching nearly one σ_p unit by t=10. Therefore, as expected, the amount of bias in the BLUP evaluation in schemes ignoring the genotype information clearly depends on the size of the QTL. These schemes led to a negative bias in the trait under complete polygenic inheritance, but only when QTL explained 50% of the genetic variance, whereas positive bias was observed for the trait under mixed inheritance for both θ levels.

5.3.5. Gamete phase disequilibrium between polygenes and QTL

The amount of disequilibrium in t_1 and t_2 for the two levels of θ in those generations of selection before fixation of the favourable allele (i.e., segregating in at least 200 replicates) is shown in Figure 5.5. Selection on an index including two traits, one affected by a sizeable QTL induced disequilibrium in both traits. For both levels of θ the scale of gamete phase disequilibrium between QTL alleles and polygenes was much greater for the trait with the greatest h^2 in the breeding goal (t_1) (i.e., see the scales of the ordinate for each trait in Figure 5.5).

For $\theta = 0.1$, the disequilibrium in t_1 in both T_G and O_G schemes was around $0.10\sigma_p$ such that gametes containing the unfavourable allele (i.e., B) had higher polygenic value than gametes containing the favourable allele (i.e., A). These schemes also created disequilibrium between polygenes and the QTL alleles in t_2, but to a much lower extent and of opposite sign (from -0.02 to -$0.03\sigma_p$). Schemes ignoring the QTL information did not generate a significant amount of disequilibrium in either trait.

For $\theta = 0.5$, the magnitude of disequilibrium in schemes using the QTL was much greater than for $\theta = 0.1$. Also, significant disequilibrium was created in t_1 not only in G schemes, but also in those schemes ignoring the QTL information (from about 0.025 to $0.050\sigma_p$).

These results indicate that although the QTL affected only one of the traits in the breeding objective, gametic phase disequilibrium between polygenes and the QTL alleles was generated for both traits, such as the polygenic mean of gametes carrying the unfavourable allele was greater. The sign of the disequilibrium differed for both traits following the sign of the polygenic response (i.e., positive for t_1 and negative for t_2).

5.4. Discussion

This study has evaluated the combined benefits of using QTL genotype information and quadratic optimisation for maximising gain while constraining the rate of inbreeding to a desired value when selection is applied on a selection index including two negatively correlated traits. This represents an extension of the work of Villanueva *et al.* (1999) based on a single trait scenario. This is relevant from a practical point of view, since in breeding schemes more than one trait contributes to the overall economic benefit. The study focused on a practical scenario which might

benefit from the use of QTL information: a trait with high heritability affected by polygenes (e.g. a production trait) negatively correlated to a trait with low heritability for which a QTL has been identified (e.g. a fitness-related trait). The evaluation included scenarios where phenotypes for the trait affected by the QTL were only available in one sex.

The use of a quadratic index for optimising the contribution of selection candidates gave similar benefits over truncation selection for improving the aggregate genotype than those found previously in single trait scenarios under mixed inheritance models (Villanueva *et al.*, 1999 and 2002b). Greatest gains in the breeding goal from quadratic optimisation over truncation selection were observed from the first generation of selection (see Figure 5.1). This was expected as at the same ΔF, quadratic optimisation allocates greater contributions to the next generation to individuals with greater selective advantage (see Chapter 3).

Optimal selection achieved higher absolute polygenic gain than truncation selection in both traits. However, the polygenic negative correlation between the traits determined its sign, which was positive in t_1 and negative in t_2. While the maximisation of genetic gains is an intrinsic property of quadratic optimisation, for this set up with two negatively correlated traits, the extra polygenic gains in the negative direction in the trait with lower heritability (t_2) might be seen as undesirable. This would be the case if t_2 was related to reproductive success or to other component of overall fitness. The only benefit of quadratic optimisation with respect to truncation selection for the trait under QTL effect arose from a faster fixation of the favourable allele when QTL information was used. However, the faster the allele was fixed, the earlier the polygenic gains in the undesired direction in t_2 started to become evident. In contrast, when using the QTL information, truncation selection was equally effective in fixing the favourable allele, and its sub-optimality from the point of view of the management of contributions avoided more dramatic negative responses in t_2.

The use of the QTL genotypes allowed increased gains in the aggregate breeding value across all generations of selection. Therefore, for *H*, the previously reported conflict between the short- and long-term gains when information from identified genes is used (e.g., Gibson, 1994; Pong-Wong and Woolliams, 1998; Dekkers and van Arendonk, 1998; Villanueva *et al.*, 1999) was not observed. This was also the case for the scenario with $\theta = 0.5$ and *full-data* even after allowing for extra 10 generations of selection to enable the long-term comparisons being held after the favourable allele was fixed in *I* schemes (data not shown).

The effect of using the QTL information on the polygenic gain of each of the traits in the breeding goal appeared to be more related to the trait heritability (i.e., the relative weight in the breeding goal) than to the particular inheritance model. The use of QTL information had a greater effect on the trait with the highest heritability (trait 1) for both levels of θ. Significant amounts of disequilibrium between QTL alleles and polygenes were found for this trait (which had a complete polygenic inheritance) in that it reduced its polygenic gain before fixation. The effect of disequilibrium in reducing polygenic responses has been reported in single trait models under mixed inheritance (e.g., Pong-Wong and Woolliams, 1998; Dekkers and van Arendonk, 1998; Schulman *et al.*, 1999), where the selected trait was affected by a QTL and polygenes. The amount of disequilibrium for t_1 was particularly important for $\theta = 0.5$ where during segregation of the favourable allele, the polygenic mean of gametes carrying the unfavourable allele was up to $0.25\sigma_P$ units greater than the corresponding polygenic mean for gametes carrying the favourable allele (Figure 5.5). This explains the lower (polygenic) response in t_1 observed of *G* schemes before fixation in both optimal and truncation selection. After fixation in *G* schemes, these schemes behaved as in a complete polygenic model and gave greater gains than *I* schemes. Also, *G* schemes were also free from any bias on t_1, whereas *I* schemes were subject to the effects of bias in the BLUP evaluation in t_1 throughout the whole selection process for $\theta = 0.5$ (Figure 5.4).

For the trait under QTL control (i.e., the trait with the lowest heritability), the disequilibrium were much lower (i.e., the polygenic mean of gametes carrying the

unfavourable allele was up to $0.04\sigma_P$ units greater than the corresponding polygenic mean for gametes carrying the favourable allele) and *G* schemes achieved greater gains across all generations than the corresponding *I* schemes, hence indicating that polygenic gains was not compromised by the use of the QTL information (even for $\theta = 0.5$). The main differences among *G* and *I* schemes in polygenic gain for t_2 arose from the bias in the BLUP evaluation in *I* schemes that was particularly important for $\theta = 0.5$.

The most important benefit from using the QTL information arose from counteracting the effects of the negative polygenic correlation, allowing for positive total gains in t_2 while the QTL was still segregating (Figure 5.3). In particular, for a gene explaining 50% of the total genetic variance in t_2, the total gain was positive for the ten generations of selection analysed, both with truncation and optimised selection. After fixation, the total gain in t_2 was solely determined by the polygenic gain and its direction was determined by the sign of the polygenic correlation (i.e., negative), which explains the decreasing total gain in t_2 in later generations. The differences in polygenic gain in t_2 between truncation schemes and between optimal schemes using or ignoring the QTL, respectively for $\theta = 0.1$ or $\theta = 0.5$ are mainly explained by the bias in the BLUP evaluation generated in T_I and O_I schemes. Whilst both T_G and O_G achieved similar corresponding polygenic gains in t_2 at both level of θ, the important amounts of positive bias generated in both T_I and O_I schemes (up to $0.97\sigma_P$) prevented greater (but negative) gains for $\theta = 0.5$ (see Figure 5.5).

The effects of bias in the BLUP evaluation in schemes ignoring the QTL information were particularly important for the trait affected by the QTL. This was expected, as in these schemes the phenotypes for this trait were uncorrected for the effects of the major gene. The bias created significant differences in polygenic gain in t_2 between *G* schemes and *I* only for $\theta = 0.5$. This can be seen as a 'beneficial' effect, since avoided extra negative gains particularly in schemes under quadratic optimisation. A significant downward bias in the BLUP evaluation was also observed in the trait under polygenic inheritance, but only for a large QTL. In this case, the bias was created from using a biased covariance in the genetic variance-covariance matrix

between the two traits. For $\rho = -0.5$, the genetic covariance between both traits used in the multi-trait BLUP evaluation when the QTL information was accounted for was –0.11 (i.e., –0.5x0.55x0.39), whereas when the QTL was ignored and $\theta = 0.5$, the genetic covariance used was –0.21 (i.e., –0.5x0.55x(0.39+0.40)). As mentioned before, this bias prevented *I* schemes from maintaining the advantages over *G* schemes in polygenic gain in t_1 after fixation of the favourable allele in *G* schemes for $\theta = 0.5$.

The greater gain in the aggregate genotype from the use of QTL information when phenotypes were available only for females in the trait under QTL effect, follows the observation of previous studies under single trait scenarios and MAS selection (e.g., Ruane and Colleau, 1996), and general expectations (Goddard and Hayes, 2002; Goddard, 2003).

The loss of the favourable allele was negligible in *G* schemes even when the QTL explained only 10% of the total genetic variance in t_2. With MAS truncation selection on an index including two negatively correlated traits, Verrier (2001) found similar allele loss both in schemes using and ignoring the QTL for $\theta = 0.1$. MAS was only able to avoid allele loss and increase its frequency when the QTL explained 20% of the total genetic variance and when the marker was very close to the QTL (i.e., a recombination rate of 0.02). Even in this case the benefits in avoiding the loss of the favourable allele from MAS with respect to conventional BLUP selection were only moderated (i.e., the beneficial allele was lost in 37% and 53% of replicates, respectively). This loss is higher than the loss found here with GAS where the favourable allele was lost in 0.4% and 33.6% of the replicates for T_G and T_I, respectively (Figure 5.1). However, the higher loss found by Verrier (2001) was also due to the fact that in his case the trait affected by the QTL not only had the lowest heritability, but also had a lower relative weight in the breeding goal (a quarter of the weight given to the trait under polygenic inheritance).

A greater loss of the favourable allele in the O_I scheme than in the T_I was observed for both levels of θ. This effect is likely to be related to the lower relative weight of

the trait under QTL effect in the breeding goal (i.e., lower h^2). Quadratic optimisation will allocate greater contributions to individuals with higher index values and the ranking of candidates will be dominated by the trait with higher h^2, that is, the trait under complete polygenic inheritance (t_1). Also, as showed previously, greater disequilibrium was created for this trait such as gametes carrying the favourable allele had a lower polygenic mean in t_1. In consequence, until fixation candidates with genotypes carrying the favourable allele will have a lower selective advantage compared to candidates carrying the unfavourable allele. By including the QTL information in the selection decisions the extra allele loss in optimal schemes disappears.

In this study it was assumed that the weight given to the QTL and the polygenic components in the total EBV were equal. There is existing framework for optimising the emphasis given to the QTL in the selection criterion for maximising the cumulated genetic gain over a planning horizon (Dekkers and van Arendonk, 1998; Manfredi *et al.,* 1998; Li *et al.*, 2001) and these optimal weights can be combined with the dynamic optimal tool for maximising gain while constraining inbreeding (Villanueva *et al.*, 2002a). This combined optimisation could be readily applied to breeding goals with multiple traits and a QTL segregating in one of the traits. Benefits from optimal QTL weights are expected in reducing the polygenic loss in the trait with greater weight in the breeding goal.

The genetic gains in the breeding goal for each of the schemes compared in this study might be considered as upper limits as, in practice, markers rather than identified genes are more likely to be used. Verrier (2001) found only small and non-significant benefits (of around 3%) from MAS over traditional BLUP truncation selection when selecting on a breeding goal with two negatively correlated traits and a QTL affecting only the trait with lower relative weight ($\theta = 0.1$ or 0.2). Here, the corresponding advantage of truncation GAS over truncation selection ignoring the QTL was 6%. In a single trait scenario, Villanueva *et al.* (2002b) showed limitations for standard BLUP MAS for approaching the gains of GAS, even for a recombination rate as small as 0.05. For MAS being able to approach the upper

limits of GAS, extra independent information on the QTL effects of selection candidates was required. They emphasised the relevance of locating the QTL for capitalising the potential benefits from including a major gene in the selection decisions. The same principle would apply in multiple-trait breeding goals.

This study has analysed the consequences in the overall breeding goal and in individual traits from selecting for identified causative mutations affecting traits with quantitative variation (e.g., Grisart *et al.*, 2002; Wiener *et al.*, 2002). This helps to address relevant questions in practical breeding schemes of how to deal with identified genes in the context of the overall economic profit. Whilst the study has highlighted the value of quadratic optimisation and the use of QTL information in such multi-trait breeding objectives, specific results are limited to the genetic models assumed and the set of parameters chosen. Nevertheless, benefits from quadratic optimisation are also expected for alternative models. Extensions of the current model to allow, for allelic interactions between QTL alleles (e.g., recessive, dominant or overdominant), for including emerging pleiotropic effects across the breeding objective (e.g., Short *et al.*, 2002; Freyer *et al.*, 2003) or to accommodate multiple QTL are straightforward.

Table 5.2. Total accumulated genetic gain in the breeding goal over generations (t) for truncation and optimal BLUP schemes selection ignoring (T_I, O_I) or using (T_G, O_G) the QTL information for two levels of the proportion of the total genetic variance explained by the QTL affecting trait 2 (θ). Phenotypic records were available for both sexes in trait 1 and only for females in trait 2. [a]

	$\theta = 0.1$				$\theta = 0.5$			
t	T_I	T_G	O_I	O_G	T_I	T_G	O_I	O_G
1	0.336	0.372	0.529	0.598	0.372	0.614	0.595	1.023
2	0.654	0.736	0.868	0.978	0.769	1.281	0.995	1.641
3	0.943	1.065	1.191	1.345	1.133	1.659	1.369	1.955
4	1.212	1.372	1.501	1.655	1.489	1.957	1.744	2.254
5	1.474	1.656	1.794	1.946	1.833	2.226	2.101	2.543
6	1.727	1.922	2.056	2.210	2.137	2.481	2.431	2.806
7	1.974	2.173	2.323	2.467	2.431	2.731	2.736	3.069
8	2.200	2.404	2.567	2.704	2.687	2.954	3.027	3.308
9	2.420	2.635	2.802	2.935	2.926	3.179	3.299	3.542
10	2.631	2.851	3.029	3.161	3.139	3.390	3.551	3.768

[a] Standard errors ranged from 0.002 to 0.020.

5.6. Figures

Figure 5.1. Frequency of the favourable allele over generations for truncation and optimal BLUP selection schemes ignoring (T_I, O_I) or using (T_G, O_G) the QTL information for two levels of the total genetic variance explained by the QTL affecting trait 2 (θ). Two scenarios are considered: 1) phenotypes are available for both traits in both sexes (*Full-data*); and 2) phenotypes for trait 2 are available only for females (*Sex-limited*). The percentage of replicates (% lost) where the favourable allele was lost is indicated. Standard errors ranged from 0.0 to 0.01.

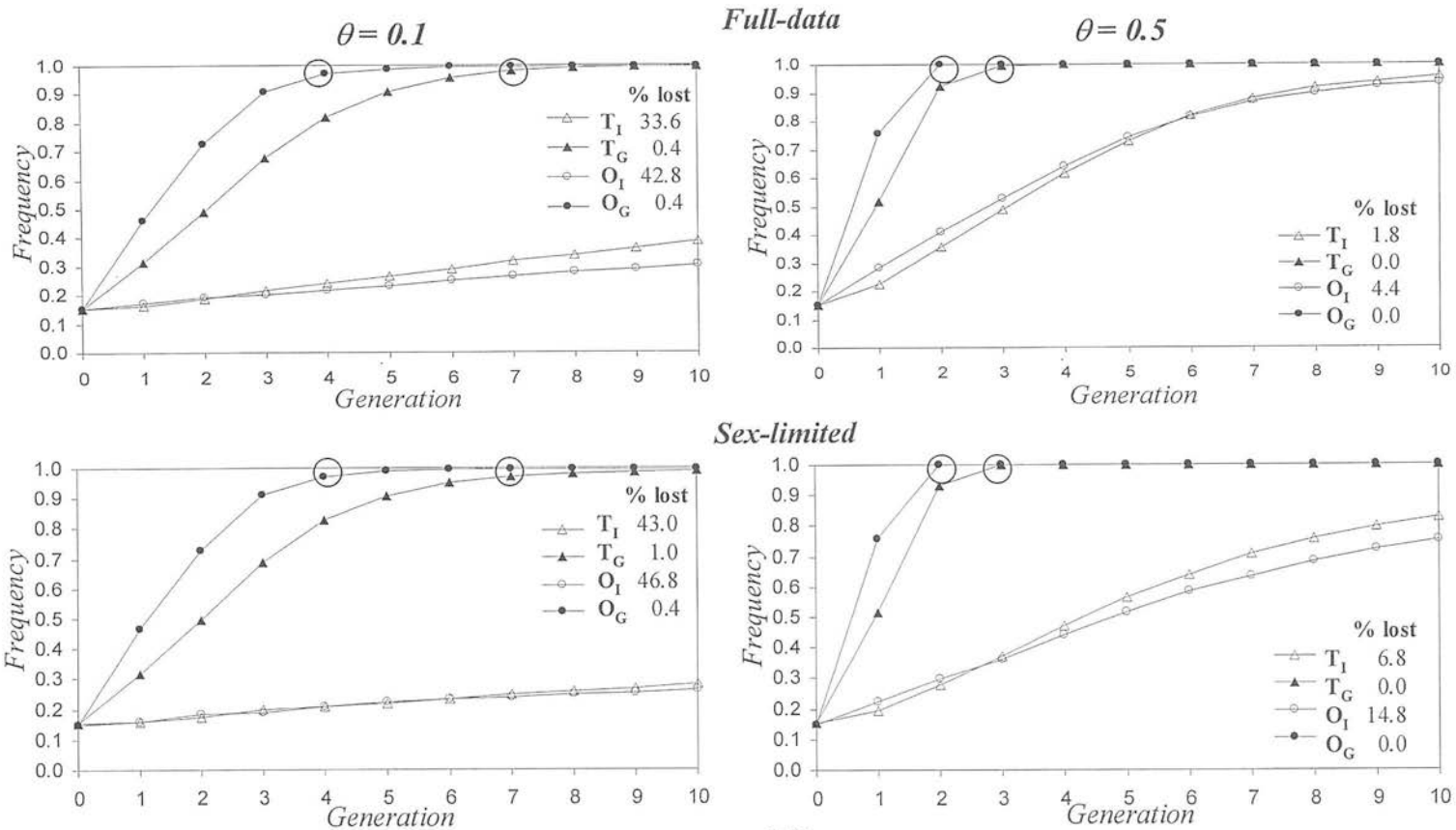


Figure 5.2. Accumulated total (polygenic) genetic gain in trait 1 over generations for truncation and optimal BLUP selection schemes ignoring (T_I, O_I) or using (T_G, O_G) the QTL information for two levels of the total genetic variance explained by the QTL affecting trait 2 (θ). Phenotypic records were available for both traits in both sexes. Standard errors ranged from 0.0 to 0.02

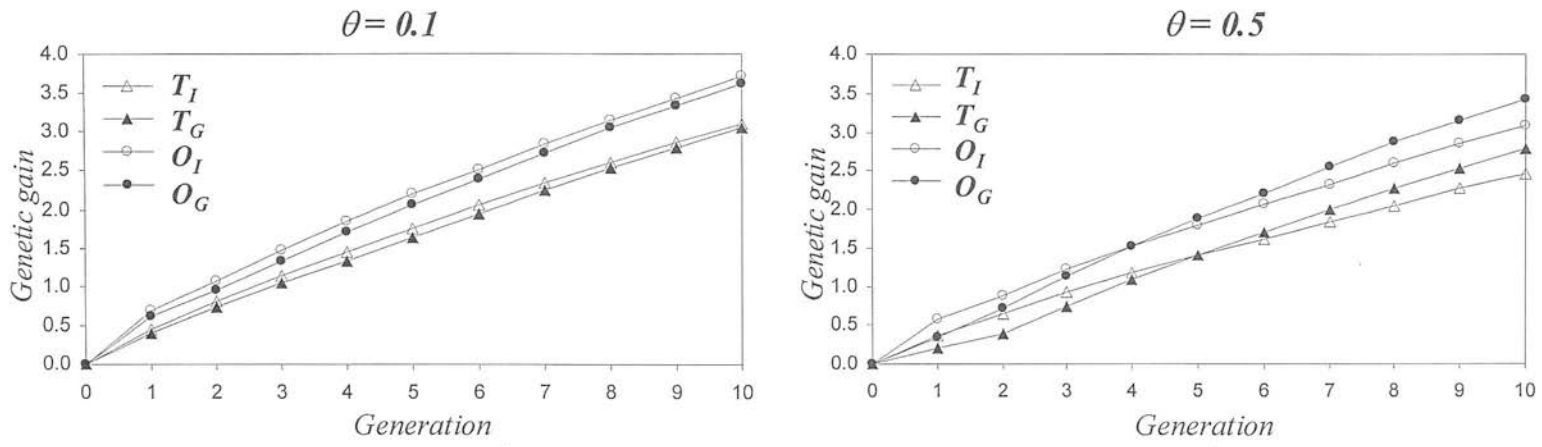


Figure 5.3. Accumulated total and polygenic genetic gain in trait 2 over generations for truncation and optimal BLUP selection schemes ignoring (T_I, O_I) or using (T_G, O_G) the QTL information for two levels of the total genetic variance explained by the QTL affecting trait 2 (θ). Phenotypic records were available for both traits in both sexes. Standard errors ranged from 0.0 to 0.01

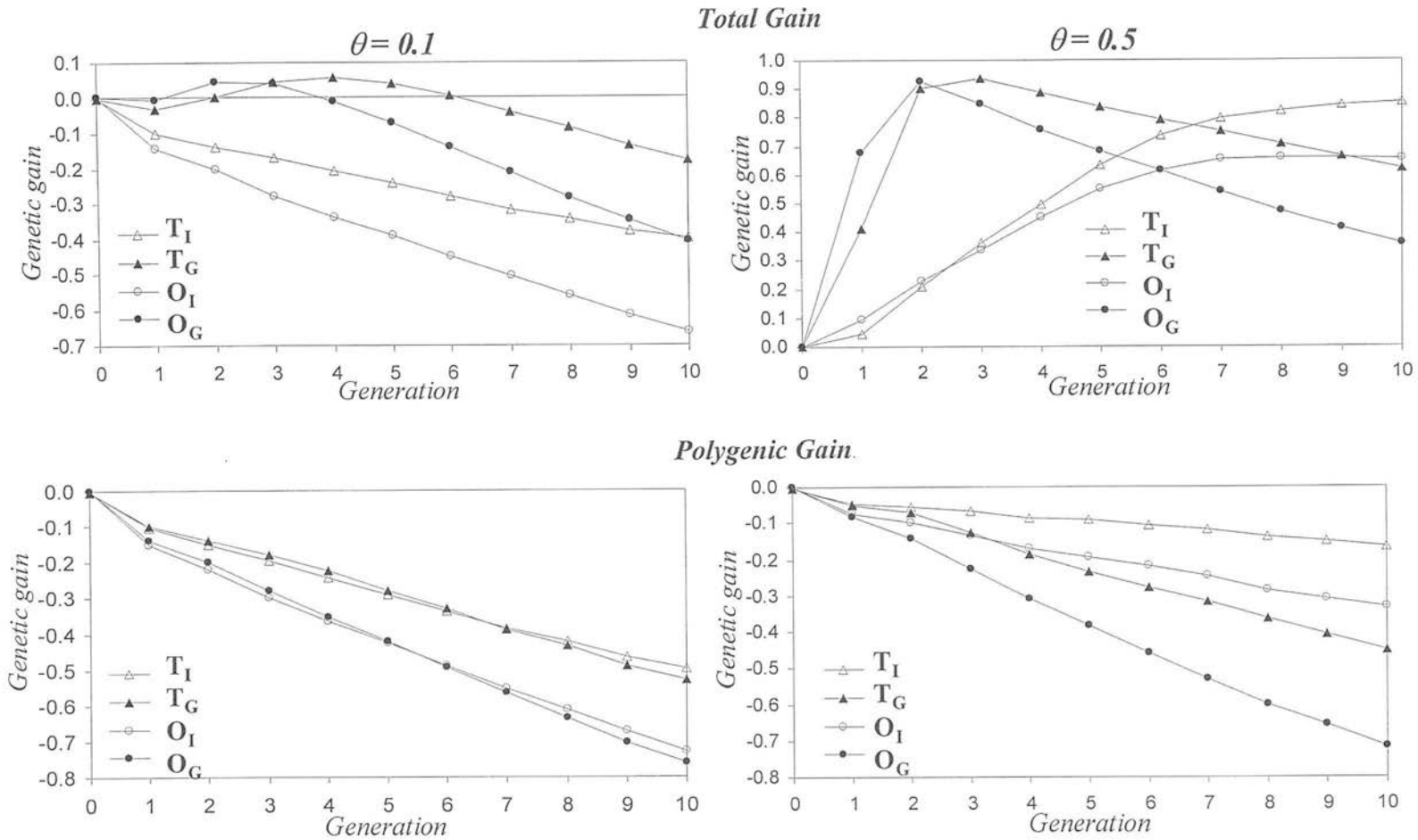


Figure 5.4. Bias (in phenotypic standard deviation units) in BLUP evaluation in trait 1 and trait 2 over generations for truncation and optimal selection schemes ignoring (T_I, O_I) or using (T_G, O_G) the QTL information for two levels of the total genetic variance explained by the QTL affecting trait 2 (θ). Phenotypic records were available for both traits in both sexes. Standard errors were 0.0 and ranged from 0.0 to 0.01 for schemes using and ignoring the QTL information, respectively.

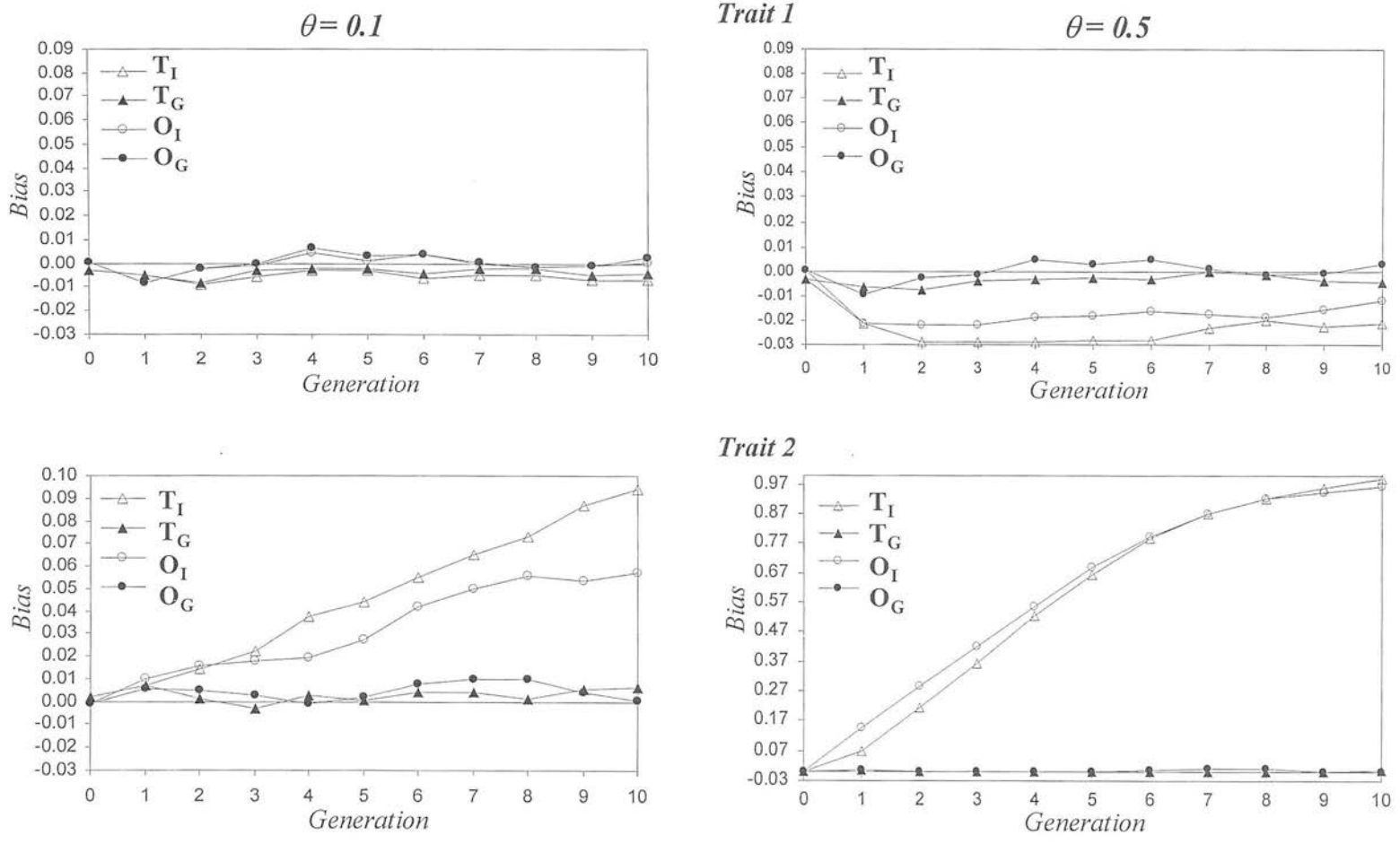
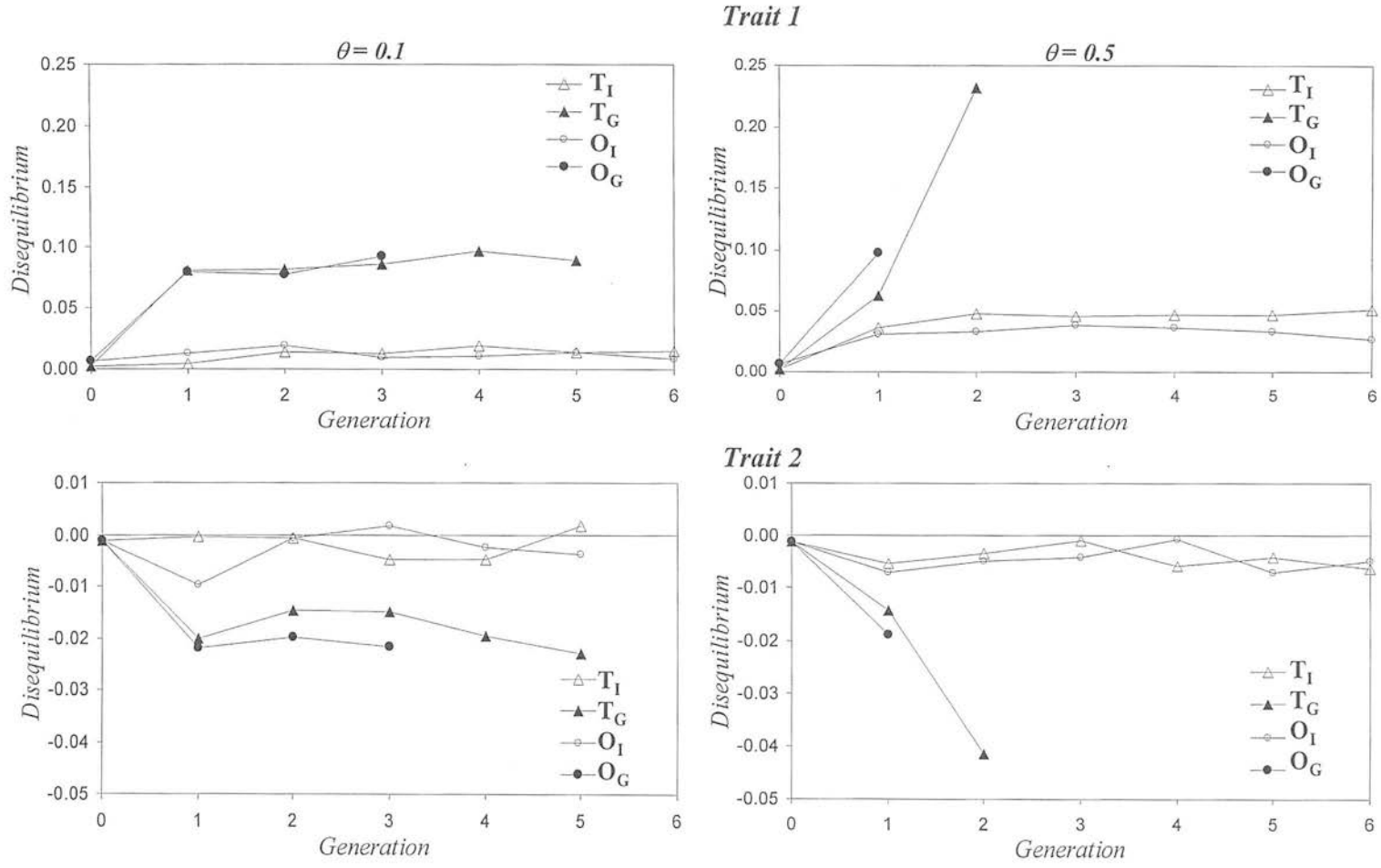


Figure 5.5. Gamete phase disequilibrium (in phenotypic standard deviation units) between QTL alleles and polygenes for trait 1 and trait 2 over generations for truncation and optimal selection schemes ignoring (T_I, O_I) or using (T_G, O_G) the QTL information for two levels of the total genetic variance explained by the QTL affecting trait 2 (θ). Phenotypic records were available for both traits in both sexes. Standard errors ranged from 0.0 to 0.05.



6. CHAPTER SIX

General Discussion

6.1. Overview of objectives

Quadratic indices constitute a general approach for the joint management of ΔG and ΔF in livestock populations subject to selection for commercially important traits (Woolliams *et al.*, 2002; Villanueva *et al.*, 2003). Furthermore, applications have been proposed in conservation programmes for deciding the optimum contribution of set of populations to gene banks (Eding *et al.*, 2002) and to select against disease genes (Sonesson and Meuwissen, 2003). The most relevant property of quadratic indices is that they keep separated the method for genetic evaluation (namely, BLUP) and the policy for controlling the rate of inbreeding in the population (Woolliams *et al.*, 2002). Hence, genetic gain is not compromised when a certain risk policy, summarised by the rate of inbreeding, is chosen. It is recognised that dynamic selection algorithms that implement quadratic indices in livestock breeding populations provide the maximum rate for a given rate of inbreeding (Meuwissen, 1997; Grundy *et al.*, 1998a). This is true not only when the estimated breeding value has only a polygenic component, but also when it includes both a polygenic and a QTL component (Villanueva *et al.*, 1999), particularly when optimum relative weights for each component of the aggregate breeding value are used (Villanueva *et al.*, 2002b).

On the other hand, the benefits of optimal selection over traditional truncation selection had not been thoroughly quantified. Rough estimates from simulation studies in the small number of rather limited scenarios considered suggested that benefits from optimal selection would range from 21% to 60% depending on the ΔF constraint (Meuwissen, 1997; Meuwissen and Sonesson, 1998; Grundy *et al.*, 2000). Moreover, the question of how much of these expected benefit would be realised in real livestock populations under BLUP selection was unknown.

With regard to the mechanics of the optimisation of contributions under quadratic indices little was known on the differences with respect to traditional truncation selection based on linear indices. This was not defined as a research objective from the outset, but was later perceived as a relevant requisite for allowing the

development of an appropriate framework for predicting the rate of gain under quadratic optimisation. The relevance of this study was that, for the first time, this thesis presented empirical evidence that the key selective advantage under quadratic optimisation is the Mendelian sampling term, in agreement with the theoretical definition of gain by Woolliams and Thompson (1994).

Therefore, this thesis has focused on the understanding of the magnitude of the benefits from optimal selection with constrained inbreeding in practical selected livestock breeding populations and has explored its benefits in multi-trait breeding goals under a mixed inheritance model. This was addressed from different angles, by using empirical approaches in the form of real data (Chapter 2), simulations with different inheritance models (Chapter 3 and 5) and deterministic prediction frameworks (Chapter 4).

6.2. Key findings

The main developments and findings of each of the four research chapters are reviewed below.

Chapter 2: The implementation of dynamic selection algorithms in real livestock populations has been successful and has demonstrated that quadratic indices constitute a potentially highly effective way of managing gain and inbreeding for a broad range of schemes in terms of scale and current inbreeding level. Substantial expected increases in ΔG compared to the current observed rates of gain were found in both Aberdeen Angus (AA) and Meatlinc (ML) populations at the corresponding observed ΔF in the last generation (0.002 for AA and 0.01 for ML). An upper bound expected benefit of six and four times the observed index gain for AA and ML, respectively was obtained when contributions of both male and females were optimised. When only male contributions were optimised conditional to a fixed female contribution equivalent to one mating (i.e.,

no selection on females was allowed) expected increases in ΔG were threefold and twofold for AA and ML, respectively. The maximum contribution allocated to a selected male candidate was equivalent to 83 and 298 matings for ML and AA, respectively. These were below maximum numbers of offspring per male observed in each population, hence, these results are realistic from an implementation point of view. When benefits from optimal selection were compared to the gain that would be achievable under traditional truncation selection based exclusively on BLUP-EBV at the observed ΔF in each population, benefits were 30% and 17% for AA and ML, respectively. These two latest figures represent the advantage from changing from traditional ways of selection based on ranking individuals and allocating equal contributions to them, to optimal selection where higher contributions are allocated to candidates with higher EBV while the desired ΔF is achieved. An operational tool has been developed which is ready to be implemented in beef cattle and sheep populations, for the routine and periodic evaluation of selection decisions in commercial environments in the UK. This operational tool is capable of dealing with several thousands of candidates (e.g., rather than hundreds typically used in simulation studies). The algorithms are general as they could be applied in other species such as dairy cattle, and indeed in species with less restricted upper limits in female reproductive rates such as poultry, fish, crustacea (e.g., shrimp) and trees. In these latter cases, the benefits from optimising both male and female contributions could be an achievable target rather than an upper bound limit as in cattle and sheep populations.

Chapter 3: This simulation study demonstrated that the dynamic management of contributions in quadratic optimisation is done with respect to the best estimate of the Mendelian sampling term. That is, under quadratic optimisation, at all times (i.e., from selection time to the convergence of the long-term contributions) the selective advantage is better explained by the estimated Mendelian sampling term, not the EBV. Therefore, in

agreement with the definition of genetic gain of Woolliams and Thompson (1994), the contribution of selected candidates to future generations is decided upon the unique superiority or inferiority with respect to the parental mean. In contrast, under traditional truncation selection on breeding values, the family average will affect whether an individual is selected or not.

The multivariate analysis of contributions (initial or long-term) indicated that the partial regression coefficient on the estimated Mendelian sampling term at convergence were much greater than those involving the ultimate estimate of the breeding value. Moreover, the contribution of EBV on decreasing the residual mean squares of the regression was always negligible. This was true for the whole set of parameters studied. In contrast, under truncation selection such clear distinction between selective advantages was not observed. Under quadratic optimisation a selected individual will have an equal or greater estimated Mendelian sampling term than that of an unselected candidate with a higher probability than under truncation selection. The opposite was true for the estimated breeding value under truncation selection.

Another relevant finding was that the optimisation outcome at the time of selection obtained from initial estimates of contributions and Mendelian sampling terms (i.e., the only information available in practice) was a good estimator of the ultimate relationship between long-term contributions and the estimates of the Mendelian sampling term at convergence. This was true for the whole range of ΔF and heritabilities studied.

A simple and intuitive measure of the efficiency of the breeding scheme for the optimum exploitation of the genetic variation as a function of the

ΔF constraint is proposed as $1-\frac{\frac{1}{4}\sum r_{dev}^2}{\Delta F}$ where $\sum r_{dev}^2$ is the sum of squared deviations of observed long-term contributions from their expected values (i.e., from linear allocation). A greater efficiency is expected as the ΔF constraint is tighter, which was confirmed from simulations (e.g., 0.92 for ΔF=0.01 and 0.84 for ΔF=0.02). Benchmark values of 0.92 and 0.50 for quadratic optimisation and truncation selection, respectively for ΔF=0.01 and h^2=0.25 were found.

Since the initial contribution of selected candidates is decided upon the available estimate on the Mendelian sampling term, the accuracy of its estimate appeared as a key central parameter for accounting for the degree of departure from the ideal linear relationship between contributions and true Mendelian sampling term (e.g., Grundy *et al.*, 1998a; Woolliams *et al.*, 2002). The need for accurate predictions of the ultimate accuracy of the Mendelian sampling term to enable deterministic predictions of the rate of gain with restrictions in the rate of inbreeding was identified.

Chapter 4: A deterministic prediction of the rate of genetic gain while the rate of inbreeding is constrained to pre-defined levels was developed. The rate of gain from quadratic optimisation was obtained as $\Delta G_{quad} = \rho_{conv} \times \Delta G_{ideal}$, where ρ_{conv} is the accuracy of the Mendelian sampling term at the convergence of the long-term contributions of selected candidates and ΔG_{ideal} is the theoretical ideal rate of gain for a given rate of inbreeding after an exact allocation of long-term contributions and Mendelian sampling terms (Grundy *et al.*, 1998a). Hence an adjustment of the theoretical prediction of the rate of gain for constrained rate of inbreeding of Grundy *et al.* (1998a) was the route followed. The prediction is self-contained, requiring only three inputs: the size of the scheme (i.e., number of candidates per generation, T), the

trait h^2 and the desired ΔF constraint. The prediction for ΔG is obtained in two steps: i. the ΔG_{ideal} is predicted with the above inputs using the formulae of Grundy *et al.* (1998a); ii. the ρ_{conv} is predicted from an extended pseudo-BLUP method (Wray and Hill, 1989) allowing for terms related to the Mendelian sampling term. Good predictions for ρ_{conv} were obtained. The number of sires and dams are required as inputs for the index. Therefore, a prediction of the effective number of parents of equal contributions (N_c) that would be obtained from quadratic optimisation at the time of selection based on 1-h^2 and the joint factor $T\Delta F$ was developed.

Predictions of ΔG_{quad} were reasonably accurate for the whole range of h^2 (i.e., excluding the upper limit h^2=0.99), particularly for small scheme sizes and tight ΔF constraints. For T=100 and ΔF=0.01 the gain was under predicted in only about -6%. For ΔF=0.025 the gain was over predicted on average in 3.4% up to h^2=0.4 and under predicted on average in -2.4% from h^2=0.5. For T=300 and ΔF=0.01 the gain was over predicted on average in 3.9% up to h^2=0.6 and under predicted on average in -3.7% from h^2=0.6, while for ΔF=0.025 the gain was overestimated on average on 11.8% up to h^2=0.8.

In this chapter, for the first time, deterministic comparisons of the rate of gain from quadratic optimisation and that from truncation selection have been performed. Results suggest that greater benefits from quadratic optimisation over BLUP truncation selection are expected for larger scheme sizes. Benchmark values for the maximum ratio $\Delta G_{quad} / \Delta G_{tru}$ for ΔF= 0.01 and 0.025 were respectively 1.40 and 1.45 for T=1,000 and 1.27 and 1.32 for T=100. It is acknowledged that these benefits may include positive effects from some degree of factorial matings in quadratic optimisation over hierarchical matings, but this would be less important as the population size increases (Sonesson and Meuwissen, 2000). Also, greater gains might be obtainable from BLUP truncation if

the weights in the pseudo-BLUP index were optimised (Villanueva and Woolliams, 1997). Nevertheless, results are clearly indicative of the potential benefits from quadratic optimisation over selection on traditional truncation selection. Relevantly, an important gap in prediction of genetic gain has been filled providing a solution for $E(\Delta G \mid \Delta F)$ given the amount of resources, the trait h^2, the equivalent number of parents at the time of selection (N_c) and the target ΔF.

Chapter 5: The aim of this study was to investigate the potential benefits from quadratic optimisation in a two-trait breeding goal when an identified QTL explains a sizeable proportion of the genetic variance in one of the traits. Index values of candidates were used for the optimisation of contributions rather than the EBV of a single trait. From the point of view of the breeding objective extra gains were observed throughout the whole selection process (ten generations) when quadratic optimisation and the use of QTL information were combined. In contrast, the selection scheme that combined both optimal selection and the use of QTL information was not best for each of the traits in the breeding objective. Optimal selection allowed extra responses in the polygenic component of the breeding objective, although the direction of the polygenic gain in each trait was a function of the negative polygenic genetic correlation and the heritability (i.e., the relative weight) of the trait. At all times, the extra polygenic gains from optimal schemes over that from truncation selection schemes were positive for the trait under complete polygenic inheritance (higher h^2), and negative for the trait under mixed inheritance (lower h^2). The use of the QTL information imposed a cost in the polygenic gain of the trait under polygenic inheritance during segregation of the favourable allele, although did not compromise polygenic gains in the long-term. The extent of the gamete phase disequilibrium was much higher for the trait under polygenic inheritance, which had the highest h^2. Before fixation, the polygenic mean of gametes carrying the unfavourable allele was greater than the

corresponding polygenic mean of gametes carrying the favourable allele. This indicates that the effects of selecting for the QTL on each trait depend on its relative weight in the breeding objective and not exclusively the inheritance model.

The use of the QTL clearly allowed not only positive gains in the trait with lower h^2 in the index, but also avoided the loss of the favourable allele. In contrast, optimal selection ignoring the genotype information led to greater loss of the favourable allele than truncation selection. This was related to the lower heritability (i.e., lower relative weight) of the trait under QTL effect in the breeding objective.

The scenario studied represents a likely practical situation in which the breeding objective includes a trait with high heritability affected by polygenes (e.g., a production trait) and a negatively correlated trait with low heritability (e.g., fitness-related trait) for which a QTL has been identified. Specific results are difficult to generalise as they are restricted to the genetic model assumed and the set of parameters chosen. Nevertheless, the interaction between multiple factors such us the size of the QTL effect, the selection method, the use of QTL information, the bias in the BLUP evaluation, the gametic phase disequilibrium between QTL alleles and polygenes has been highlighted. This provided important insight on the consequences on the overall breeding objective and each of the individual traits. Moreover, it has been shown that benefits from quadratic optimisation through the optimisation of contributions are expected when selection is on multi-trait selection breeding goals for the maximisation of overall economic profit. This generalises the potential benefits from the use of quadratic optimisation already observed when selecting on single traits.

6.3. Implications

6.3.1. Expected benefits from quadratic optimisation over BLUP truncation selection

The benefits of quadratic optimisation over truncation selection obtained in this thesis from empirical data (real data or stochastic simulation) and from predictions of ΔG_{quad} (using the extended pseudo-BLUP method developed in Chapter 4) and ΔG_{tru} (using the program Sel*A*ction of Bijma and Rutten, 2002) are summarised in Table 6.1.

Table 6.1. Observed and predicted ratios of the rate of gain from quadratic optimisation on the rate of gain under BLUP truncation selection ($\Delta G_{quad} / \Delta G_{tru}$) for a range of inbreeding rates (ΔF, %), heritabilities (h^2) and numbers of candidates per generation (T). Observed $\Delta G_{quad} / \Delta G_{tru}$ are taken from studies in this thesis on real livestock populations (Aberdeen Angus and Meatlinc) and simulated data under polygenic and mixed inheritance models. Predicted $\Delta G_{quad} / \Delta G_{tru}$ are calculated using population parameters from the empirical studies (real or simulated) in this thesis using the approach in Chapter 4 for predicting ΔG_{quad} and Sel*A*ction for predicting ΔG_{tru}.

Source	Scenario		Parameters			$\Delta G_{quad} / \Delta G_{tru}$	
			ΔF	h^2	T	Observed	Predicted
[1] Real data	Aberdeen Angus		0.2		6,400	1.30	1.40
	Meatlinc		1.0		1,300	1.17	1.38
Simulated	Polygenic						
	[2] Single-trait		1.0	0.25	100	1.13	1.12
	Mixed inheritance						
	[3] Two-trait		6.0	0.29	120		1.71
	$\theta = 0.1$	I				1.04	
		G				1.06	
	$\theta = 0.5$	I				1.12	
		G				1.03	

[1] Observed ratios from Chapter 2; Table 2.2.

Predicted ratios were obtained after using a range of heritabilities from 0.30 to 0.50 (Jones, personal communication) for both the Beef Index (beef cattle) and the Lean Index (sheep). Predictions were not sensitive to changes in heritability.

[2] From Chapter 3; ΔG from generation 3 to generation 4 (i.e., referred to as 'selection time')

[3] From Chapter 5 Table 5.1; ΔG the breeding goal from generation 9 to generation 10. The symbol θ represents the proportion of the total additive genetic variance explained by the identified QTL affecting the second trait in the breeding objective. The symbols I and G represent schemes either ignoring or using genotype information in the BLUP evaluation. The h^2 corresponds to the value of a two-trait index selection under a polygenic model obtained from Sel*A*ction (i.e., h_1^2=0.3 and h_2^2=0.1, and $\sigma_{P_1}^2 = \sigma_{P_2}^2 = 1.0$ for trait 1 and trait 2, respectively, and a genetic correlation of –0.5)

The observed benefits in ΔG from quadratic optimisation over truncation selection at the same ΔF ranged from 13% to 30% when real or simulated data for a single trait under a polygenic model were used. The corresponding deterministic predictions of benefits were very accurate for the case of simulated data. In contrast, for real livestock populations, the predicted benefits were higher than the observed. The main reason for this overprediction was that the deterministic prediction of ΔG_{quad} does no takes into account any limitation in reproductive capacity in each sex. The observed benefits from quadratic optimisation followed the general trend presented in Chapter 4 (i.e., predicted $\Delta G_{quad} / \Delta G_{tru}$ increases as the size of the scheme increases; see Figure 4.5).

Surprisingly, similar benefits from quadratic optimisation (around 1.4) were predicted for both Aberdeen Angus and Meatlinc populations. Although both populations differed substantially in size (T), the corresponding ΔG_{ideal} (i.e., around $1.1\sigma_A$) and $T\Delta F$ (i.e., 12.8 and 13.0 for AA and ML, respectively) were very similar. Consequently, similar predictions of the ratio N_c / N_r (i.e., around 3.0) were obtained that led to similar predictions of the accuracy of the Mendelian sampling term (i.e., around 0.7) and of ΔG_{quad} (i.e., around $0.8\sigma_A$). In addition, at the corresponding ΔF similar predictions of ΔG_{tru} were obtained from Sel*A*ction (not shown).

The deterministic prediction of $\Delta G_{quad} / \Delta G_{tru}$ for the two-trait model clearly overpredicted the observed ratio. The most evident reason for the overprediction is that different underlying genetic models (i.e., polygenic for the prediction and mixed inheritance model for the empirical data) are involved. The predicted $\Delta G_{quad} / \Delta G_{tru}$ of 1.7 was obtained after constraining ΔF to 0.06, which in practice means a very low constraint. Hence, it could be argued that the prediction framework was pushed too far, giving an odd rather than a reasonable result. Further, let assume that the QTL explained all the genetic variation in the breeding goal and that the trait h^2 was 0.99. In this case, the corresponding predicted $\Delta G_{quad} / \Delta G_{tru}$ would be around 1.2 (result not shown, and see also Figure 4.1) rather than 1.7. The predicted and observed $\Delta G_{quad} / \Delta G_{tru}$ also differ in that for the prediction, the effects of inbreeding on genetic variance were not considered. The observed $\Delta G_{quad} / \Delta G_{tru}$ was taken at the last generation, when the cumulated effects of inbreeding on genetic variance were greatest. Nevertheless, this is not expected to have affected the observed ratio, as inbreeding affected the genetic variance in the same way in both optimal and truncation schemes.

The above results indicate that with the current framework, reasonable upper bounds for the benefits from quadratic optimisation over truncation selection can be predicted, provided that reasonable values for the ΔF constraint are used.

6.3.2. *Implementation of operational and design tools for quadratic optimisation in selected populations*

A complete framework for the implementation of quadratic indices in livestock breeding programs is now available. From a practical standpoint, in addition to the constraints on ΔF and the maximum contribution per sex, the operational tool can accommodate additional constraints on male and female contributions. On the other hand, the design tool assumes unlimited reproductive rates in both sexes. Nevertheless, even with this limitation, results from Table 6.1 suggest that the design

tool offers a reasonable upper bound approximation for the benefits that would be expected from quadratic optimisation.

For an ongoing breeding scheme currently operated under truncation selection, a two step straightforward implementation of the tools for maximising gain while ΔF is restricted to a pre-defined value can be envisaged:

i. Design Tool: To predict the benefits from quadratic optimisation given the size of the scheme (i.e., number of selection candidates per generation), heritability of the trait under selection (or the selection index), and constraint in the inbreeding rate (i.e., the risk policy). This constitutes a relevant step for the evaluation of the benefits of changing from the current selection policies to optimal selection at the same rate of inbreeding.

ii. Operational Tool: Implement the dynamic selection algorithm each round of selection for finding which candidates of each sex should be selected and their optimal contributions to the next generation. Benefits lower than those predicted at the design stage may arise if upper limits for realising the optimal mating proportions in one sex exist (e.g., from reproductive limits). Hence, the extra cost of lifting those constraints (e.g., by using reproductive technologies) might need to be evaluated against the loss in the maximum expected rate of gain. The Design Tool could be used again at this stage to calculate the predicted gain after imposing limits in the mating proportion of a particular group of candidates (e.g., a fixed contribution for females, or a maximum contribution per male). In this case, the 'observed' sum of squares of mating proportions (i.e., $N_c = 1/\sum c^2$) can be inputted to obtain the predicted rate of gain. Hence, the predicted loss in gain after imposing limits to the optimal contribution will be obtained.

6.3.3. Defining the appropriate constraint on ΔF when applying quadratic indices in livestock populations

The definition of the appropriate value of ΔF to be used in the quadratic optimisation is a key issue as it determines the risk policy to be implemented in the breeding scheme. There is a range of suggested values for minimum acceptable levels of ΔF depending on the genetic model assumed and on the criterion used.

Woolliams and Meuwissen (1994) found that ΔF from 0.2% to 1.6% (i.e., N_e from 250 to 31) would be low enough to achieve a balance between the decrease in fitness from inbreeding depression and the increase in fitness from natural selection when fitness and the response to artificial selection are not correlated. The critical minimum effective population size to achieve the balance was $N_e = D/2\sigma^2_{wa}$, where D is the inbreeding depression of fitness with complete inbreeding, and σ^2_{wa} is the additive variance of fitness. On the other hand, when fitness and the response to artificial selection were negatively correlated the critical N_e was much higher, and in some circumstances, a decline in fitness could not be prevented by increasing the effective population size. Woolliams and Meuwissen (1994) hence, suggested that such negative correlations between artificial and natural selection should be avoided. In fact, there is evidence of negative correlations between production and components of fitness in selected populations (e.g., Pryce *et al.*, 2002 predicted negative effects of selecting for milk yield alone on calving interval in dairy cattle).

A different approach for deciding the appropriate ΔF is that of Goddard and Smith (1990) and Goddard (1992). They maximised the economic value of genetic response while accounting for the effects of inbreeding on genetic variance and phenotypic depression. The recommended values for ΔF ranged from 0.2% to 1.3% and thus were within the range given by Woolliams and Meuwissen (1994).

In general, a rate of inbreeding of 1% per generation is usually considered as the minimum acceptable to avoid significant inbreeding depression in fitness. For

instance, ΔF of 1% has been used as a general recommendation for conserved populations (FAO, 1998). Also, Bijma (2000) suggested that ΔF around 1% would be sufficient for ensuring fixation of favourable alleles (Caballero and Santiago, 1998) with moderate selective advantage (e.g., 0.07 phenotypic standard deviations) and for maintaining reasonable levels of heritability when mutation contributes to the genetic variance. Frankham *et al.* (2002) however have suggested that ΔF should be much lower than 1% for balancing the loss of genetic variation by drift and that gained from mutation.

In some circumstances populations have managed to be successfully viable with ΔF per generation greater than critical values to avoid decline in fitness (e.g., values from Woolliams and Meuwissen, 1994). For instance, the Chillingham cattle, a closed population kept with no immigration for at least 300 years in a park in England, has shown a ΔF of 6.5% per generation with no signs of decline in fertility or viability (Hall and Hall, 1988; Visscher, *et al.*, 2001). The phenomenon of 'purging' deleterious alleles by natural selection is an explanation for the lack of decay in fitness in this and other highly inbred populations (Lacy and Ballou, 1998; Visscher *et al.*, 2001).

Relevant information for deciding the appropriate value for the constraint in ΔF would be provided from analysing the effects of inbreeding in the phenotypic expression of recorded traits in the breeding scheme. Significant levels of inbreeding depression have been documented in livestock populations (e.g., Lamberson and Thomas, 1984; Burrow, 1993; Thompson *et al.*, 2000; Rodrigañez *et al.*, 1998; Fernández *et al.*, 2002). In this case, stringent constraints giving rise to lower than the observed ΔF in the population would reduce the rate at which phenotypic means are reduced by inbreeding. Also, the increasing availability of molecular information provides information on the allele frequency of both rare deleterious mutations (e.g., Agerholm *et al.*, 2001) and beneficial mutations with effects on traits of economic importance (e.g., Grisart *et al.*, 2002; Wiener *et al.*, 2002). Predictive framework for the probability of fixation or loss of such alleles (e.g., Hill, 2000; Caballero and Santiago, 1998) can provide critical ΔF values for setting the appropriate constraint.

While it appears difficult to draw general recommendations for the appropriate levels of ΔF to avoid negative effects from inbreeding, the value for the restriction in ΔF can be objectively determined in practical breeding programmes. By setting the constraint to the observed ΔF, extra responses can be obtained compared to those obtained with current selection methods at the current level of risk. In Chapter 2 it was shown that substantial increases in ΔG are expected when restricting ΔF to the observed values of 0.2% and 1.0% per generation for Aberdeen Angus and Meatlinc respectively.

There are circumstances where livestock breeds are selected for traits of economic relevance but conservation is a primary objective. For instance, Gutiérrez *et al.* (2003) provided a thorough description of the population structure of eight contrasting Spanish autochthonous beef cattle populations, estimating ΔF ranging from 0.39% to 2.40%. The authors concluded that a single practical recommendation for all populations for maintaining genetic diversity does not exist. On the other hand, a simple general recommendation would be to implement quadratic optimisation for explicitly setting objectives in terms of genetic gain and inbreeding for each particular circumstance. The approach followed in this thesis of setting the constraint in ΔF to the value observed in the population appears to be a reasonable first step since, at least, the risk is not increased. In addition, Villanueva *et al.* (2003) derived a quadratic algorithm for minimising ΔF for a pre-defined ΔG. In this case, the objective function is designed to minimise $\mathbf{c}^{\mathrm{T}}\mathbf{A}\mathbf{c}$ while satisfying $\mathbf{c}^{\mathrm{T}}\mathbf{g} \geq \Delta G$.

Villanueva *et al.* (2003) pointed out that the difference between selection and conservation in farmed populations lies in the relative weight given to ΔF and ΔG in the optimisation rather than being separate frameworks. Hence, rather than genetic improvement and conservation being seen as separate objectives, they can be taken as the extremes of a broader optimisation problem.

When the only objective is to achieve the lowest possible ΔF (i.e., when conservation of genetic resources is the only goal), the ultimate solution comes from minimising

the sum of squares of long-term contributions of breeding individuals (i.e., $\Delta F = \frac{1}{4}\Sigma_{i=1}^{N} r_i^2$) over multiple generations (Sánchez *et al.*, 2003). Since over a generation the sum of the long-term contributions for N breeding individuals is one (Woolliams *et al.*, 1999), $E(r^2) = E(r)^2 + \sigma_r^2$ and $E(r) = 1/N$ the minimisation of ΔF problem relies on the minimisation of the variance of long-term contributions of breeding individuals (Sánchez *et al.*, 2003). While quadratic optimisation has the scope for minimising the sum of squares of initial contributions of selected individuals, it is not able to set individual contributions to the desired values over multiple generations (Woolliams *et al.*, 2002; Chapter 2). In Chapter 2 it was shown that the extent of the deviations with respect to the desired contribution (i.e., $\sum r_{dev}^2$) is a function of the ΔF constraint. Sánchez *et al.*, (2003) proposed a breeding scheme that manages contributions over multiple generations and removes all uncertainties about future contributions of offspring. They showed that the lowest ΔF for a population with different breeding males and females can be predicted from $\Delta F = \frac{4}{3}[1 + 2(\frac{1}{4})^d]/16M$ where d is the mating ratio and M the number of breeding males.

6.3.4. The selective advantage in quadratic optimisation

The demonstration that under quadratic optimisation the selective advantage is the Mendelian sampling term constituted a major breakthrough in the understanding of the nature of quadratic optimisation. It provides a link between the way in which genetic gain is achieved under quadratic optimisation and the definition of gain by Woolliams and Thompson (1994). The identification of the key role of the accuracy of the Mendelian sampling term in determining the outcome of the optimisation enabled the development of the deterministic prediction of the rate of response presented in Chapter 4.

These findings represent a new way of interpreting the process by which genetic gain is being achieved by selection. With traditional truncation selection (i.e., based on

linear indices) candidates are selected according to their EBV, hence, their contribution to the future genetic pool is linked to how good the family to which it belongs is. This means that individuals with poor Mendelian sampling term but with good parental average might by selected. In contrast, under quadratic optimisation the number of offspring and the future long-term contribution of a candidate depend primarily on its own independent and unique superiority (or inferiority). This is the underlying idea behind selection based on within family deviations, used as a method for controlling inbreeding (e.g., Toro and Perez-Enciso, 1990; Hill *et al.*, 1996) but with the key difference is that under quadratic optimisation individuals with greater family deviations (i.e., Mendelian Sampling term) contribute more, rather than equally to the next generation. The dynamic management of contributions by quadratic optimisation implies a balance between within-family selection (Dempfle, 1975) and between family selection. Dempfle (1975) showed that by selecting only the best male and female from within each family (i.e., full or half- sibs) greater long-term gains are expected, particularly when family sizes are big and the selection accuracy is high. Quadratic optimisation represents a more flexible approach where the balance between within and between family selection is managed for maximising the rate of gain. More stringent ΔF constraints lead to greater emphasis on within family selection whereas the accumulation of information on Mendelian sampling terms across generations lead to greater emphasis on between family selection.

The idea that the selective advantage of a particular individual is represented by its Mendelian sampling term was first proposed by Woolliams and Thompson (1994). They decomposed EBV in terms of individual estimated Mendelian sampling terms and proposed the management of genetic gain and inbreeding by controlling the weight given to estimated Mendelian sampling terms of ancestors. Grundy *et al.* (1998b) implemented such 'Mendelian index' (linear) but although proved to be successful for controlling inbreeding, it led to decreases in short-term gain. Thus, although the concept of using the Mendelian sampling term as the selective advantage was appropriate, the simultaneous management of gain and inbreeding is more effectively done through quadratic implementation than through linear indices.

6.4. Further developments

6.4.1. Operational tools

The development of algorithms to deal with quadratic optimisation is currently in an advanced stage. A number of dynamic selection algorithms are available for, in principle, covering most of the situations found in practical breeding schemes. On the other hand, apart from the one applied in this thesis (i.e., Meuwissen, 1997; Grundy *et al.*, 1998a), selection algorithms have been tested on simulation studies rather than in practical situations. Hence, the greatest challenge is to put in practice the available operational tools and most efforts should be concentrated on resolving problems related to the practical implementation. That is, to reduce the existing gap between the theoretical development and the realisation of expected benefits from quadratic optimisation in practical breeding schemes. Two situations of practical interest are discussed, overlapping generations and the use of mating strategies after the selection step. Also, the potential of evolutionary computation for providing global optimums for a constrained optimisation problem is discussed.

6.4.1.1. Overlapping generations

The operational tool used in Chapter 2 for the optimisation of Aberdeen Angus and Meatlinc did not take into account that in reality, generations typically overlap. This means that the restriction on ΔF was achieved without taking into account that some individuals in the set of candidates considered have already had some contributions during previous selection rounds. The single round optimisation in Chapter 2 included candidates born within a range years resembling a generation interval, hence partially accounting for the overlapping generation structure. Ideally, an operational tool to be applied routinely should also account for the previous contributions of candidates in breeding ages already selected in previous cycles.

Meuwissen and Sonesson (1998) and Grundy *et al.* (2000) extended the algorithms of Meuwissen (1997) and Grundy *et al.* (1998a) and described two alternative dynamic selection algorithms for overlapping generations. In this case, the restriction is applied on the average relationship across age classes $\mathbf{r}'\overline{\mathbf{A}}\mathbf{r}$, where $\mathbf{r}$ is a vector (q x 1) of weights for each of the q age classes and $\overline{\mathbf{A}}$ is a (q x q) matrix of the average relationship between age classes (see Meuwissen and Sonesson (1998) for details). The relative weights within different classes were calculated as the long-term contribution of the different age classes to the future gene pool using the gene flow method of Hill (1974). A key part of the optimisation was to derive the optimum contribution of each age class, which depended on the distribution of parents over age classes. Meuwissen and Sonesson (1998) used the algorithm of Meuwissen (1997) in an iterative approach to optimise both the contributions within each age class and the optimal age class distribution in a single step. In contrast, Grundy *et al.* (2000) first optimised contributions within each class using the algorithm of Meuwissen (1997) and found the optimum age structure by using a simulated annealing technique (Press *et al.*, 1992). Sonesson *et al.* (2000) thoroughly compared these approaches while constraining the ΔF per year, observing that both algorithms were optimal.. The algorithm of Grundy *et al.* (2000) although computationally more expensive for attaining the global optimum was more flexible as it allowed different weights for each age class. More relevantly, the tool of Grundy *et al.* (2000) was able to keep the ΔF constraint to the desired value and to set the constraint both on a per year or a per generation basis.

The optimisation of the age structure in these studies required a great deal of computing resources, particularly that of Grundy *et al.* (2000). Sonesson *et al.* (2000) found that the annealing optimisation approach required about 41 times more CPU time that the iterative method of Meuwissen and Sonesson (1998). These authors evaluated a maximum of 1,700 candidates in their simulations, requiring 30 minutes per each year and per replicate, finding that the CPU time increased by a factor n^3 where n is the number of times by which the scheme is increased. For instance, a rough estimate for the 7,000 candidates included in the Aberdeen Angus optimisation in Chapter 2 yields 64 times more time than that required for the 1,700 scheme. This

represents 1,920 minutes or 32 hours, which does not sound extremely demanding. In contrast, it would be unaffordable with the approach of Grundy *et al.* (2000) which would only be feasible in small-scaled populations or in small breeding nucleus (e.g., more commonly found in pigs and dairy cattle).

While the optimisation of age structures give important insight for the optimal design of the breeding program, the opportunities for changing the generation interval might be rather limited in an ongoing breeding scheme (this is certainly the case in beef cattle and sheep populations). When thinking on extending the operational tool used in Chapter 2 to account for overlapping generations, the problem can be simplified. In this case, the age structure and their corresponding weights (in the **r** vector) are already given. Thus, the operational tool of Mewissen and Sonesson (1998) could be used after including two extra inputs: the age class of each candidate and its previous contribution.

It should be stressed that the tool of Mewissen and Sonesson (1998) does not allow for setting the constraint on the desired ΔF per generation. Hence, from the point of view of management of long-term risk, the dynamic tool of Grundy *et al.* (2000), although computationally costly, is the appropriate approach and also the only available method. Thus, a trade-off arises between practicality and appropriateness of the two available approaches.

6.4.1.2. Mating systems

Throughout this thesis, mating of selected candidates resulting from the quadratic optimisation has been assumed to be at random. While selection is the most relevant step when controlling ΔF, additional genetic gains can be obtained when non-random mating strategies are combined to quadratic optimisation. Sonesson and Meuwissen (2000 and 2002) have shown both in discrete and overlapping generations that up to around 20% extra ΔG can be obtained when the optimisation of contributions is followed by a non-random mating step, particularly when population sizes are small

and constraints on ΔF are stringent. From these studies, minimum co-ancestry mating (MC) with restricted number of full-sibs per mating appears to be the best alternative to random mating. The basic principle of MC is to find the combination of matings among selected candidates that minimises their average co-ancestry, hence the average inbreeding coefficient in the offspring is minimised. Sonesson and Meuwissen (2002) point out that when a restriction in ΔF is used at the selection stage, a further reduction in ΔF is not expected from MC, but leads to an increase in selection differential from an improved family structure. The homogenisation of relationships across families reduces otherwise increased relationships among animals with higher EBV. This gives more scope for quadratic optimisation to allocate greater contributions to individuals with higher EBV. Also, the lower inbreeding coefficient in the offspring would lead to larger Mendelian sampling variance hence leading to increased gains.

To solve the minimisation problem of MC for n_s and n_d selected males and females respectively a matrix **F** (n_s x n_d) is set up where f_{ij} is the co-ancestry coefficient of the pair. Simulated annealing approaches have been used to find the combination of matings that minimises the average co-ancestries in **F** in both discrete and overlapping generations (see Sonesson and Meuwissen, 2000 and 2002 for details).

The implementation of a mating policy such as MC seems a natural and valuable step after the use of the operational tool in Chapter 2. This would be particularly valuable in the case of the Meatlinc Sire Reference Scheme where the size of the scheme is relatively small, the ΔF is relatively high and the control in the management of matings might be possible. For a scheme with about 800 candidates, Sonesson and Meuwissen (2002) achieved the minimisation of co-ancestries in 54 (when generation intervals were optimised) or 11 minutes (when generation intervals were not optimised). Thus, the implementation of MC even in schemes greater than the Meatlinc (e.g., about 1,200 candidates) does not appear to be extremely complicated.

A key element for the successful implementation of an operational tool with constrained ΔF in real livestock populations relies on how well practical constraints are included and handled. In Chapter 2, in addition to the quadratic constraint on ΔF and the linear constraint on the maximum contribution per sex, extra linear constraints for setting a fixed contribution for all female candidates and a maximum number of mates per male were implemented following Mewissen (1997, Appendix 1). Although this would enable the implementation of the operational tool in a wide number of practical schemes, more flexible constraints may be desired. For instance, some flexibility in setting maximum number of males and females having contributions above certain threshold might be desirable. In theory, these could be done deterministically using Lagrangian multipliers (λ) and adding extra terms to the objective function. However, the solution could be very difficult to obtain algebraically.

Evolutionary computation have received a lot of attention regarding their potential as optimisation techniques for complex real-world problems (see Schoenauer and Michalewicz , 1997 for an overview). These are based on stochastic optimisation rather than deterministic optimisation (e.g., Lagrangian multipliers). The basic idea behind evolutionary computation is to evolve a set of candidate solutions towards the maximisation of a 'fitness' function mimicking natural evolution and its principle of 'survival of the fittest'. As in nature, 'genetic' operators like crossover and mutation are used to create new 'offspring' solutions and selection ensures that better candidate solutions compose future generations. The evolutionary search of the global optimum is done within a searching space defined by the particular set of constraints used.

In particular, algorithms such as evolutionary and genetic algorithms have been postulated as the 'state of the art' approach for handling the most varied number of constraints in a flexible way for animal breeding applications (Kinghorn *et al.*, 2002). With regard to inbreeding, genetic algorithms have been implemented for optimising

both selection and matings in a single step, in the so called 'mate allocation' and uses a cost function for penalising increased co-ancestry (Kinghorn, 1998). In this approach, it is not clear whether the ΔF or the average inbreeding coefficient in the offspring is constrained. Also, by optimising both the selection and mating stages, it is not explicitly recognised that the fundamental step for managing inbreeding relies on deciding which candidates to use and how much they should contribute to the next generation. Indeed, effective management of inbreeding can be achieved by random mating while the average coancestry of selection candidates is restrictied, as done under quadratic optimisation. As mentioned before, non-random mating systems can further increase genetic merit through an improved family structure across the population.

Correnti (2002) developed an evolutionary computation strategy to explore the feasibility of finding a global optimum while constraining ΔF to a pre-defined value using the Meatlinc data from Chapter 2. The constraints applied were those described in Chapter 2 plus additional (more flexible) constraints on the reproductive capacity of selection candidates. The problem proved to be extremely challenging and although the desired ΔF was achieved along with all other constraints, the expected genetic gains were consistently lower than those achieved using Lagrangian multipliers (i.e., using the same set of constraints). The main difficulty was the handling of a large-scale 'chromosome' (i.e., the array containing the solution for each candidate) with about 1,300 'genes' (i.e., the optimal contribution for each candidate) during the search of the global optimum. Nevertheless, evolutionary computation constitutes a promising route with plenty of scope for the development of operational tools for the strategic optimisation of risk management (i.e., ΔF and $V(\Delta G)$) in breeding programs, particularly when finding explicit maximums by deterministic means becomes very complex.

6.4.2. Design Tools

The development of the quadratic optimisation design tool represented the most challenging objective of this thesis. Different approaches were evaluated to finally arrive to the current self-contained framework which requires only specification of h^2, T and ΔF. As it was presented in Chapter 4, the current approach requires two simple components: a deterministic prediction of ΔG_{ideal} for an exact allocation of long-term contributions and Mendelian sampling terms of selected candidates (i.e., from Grundy *et al.*, 1998a) and a deterministic prediction of the accuracy of the Mendelian sampling term (i.e., from an extended pseudo-BLUP index *a la* Wray and Hill, 1989).

Whilst the approach derived here is effective for predicting the benefits from quadratic indices, the main research challenge is to derive appropriate general algebra as it exists for truncation selection. For instance, under truncation selection using linear indices, the selection differential can be simply predicted using simple linear regression theory as $\hat{S} = b_{g,I} i \sigma_I$, where $b_{g,I}$ is the regression of the true breeding value in the selection index used (i.e., BLUP), i is the selection intensity and σ_I is the standard deviation of the selection index. In contrast, the question of how to predict the selection intensity using quadratic indices still remains unsolved and was a major limitation for the derivation of deterministic expressions for the rate of gain. To date, most of the knowledge on the mechanics of quadratic optimisation arises from empirical studies and the explicit formal expressions for the components of genetic gain are very limited. In fact, the only algebra available is for predicting ΔG_{ideal} derived by Grundy *et al.* (1998a) and for predicting the accuracy of the Mendelian sampling term derived in this thesis.

Three further developments of the current prediction approach that would be linked to the success in deriving explicit algebra for quadratic indices are identified.

First, the regression approach for predicting the ratio N_c / N_r that enabled a prediction of N_c was based on simulated data. Strictly, this represents a limitation, as the prediction would be rather limited to the parameter range of h^2, T and ΔF used. The

basic problem of predicting N_c is basically providing an '*a priori*' answer to the ultimate question of which would be the outcome of the optimisation for a combination of h^2, T and ΔF. The description of the process of going from selection to convergence summarised by N_c/N_r seemed the intuitive route to be followed, but further research is required to describe the process in a deterministic manner. This would represent a major achievement in the understanding of the intrinsic nature of the dynamic quadratic optimisation

Second, the deterministic prediction of ΔG_{quad}, is not flexible from the point of view of accounting for practical constraints as it for instance, assumes no restriction on the reproductive capacity of selection candidates. Hence, a gap exists between the operational and the design tool that in principle may limit the joint use of both for tools in livestock populations where the flexible management of restrictions in reproductive is essential (e.g., cattle and sheep). Nevertheless, results in Table 6.1 show that the available design tools allowed good approximation for predicting the benefits from quadratic optimisation in the Aberdeen Angus and Meatlinc populations.

Third, a relevant development would be to derive a prediction of the sum of squares of deviations of long-term contribution from their expected values after an exact allocation with the Mendelian sampling terms ($\sum r_{dev}^2$). This would enable a prediction of the effectiveness in the use of genetic variation under quadratic optimisation by alternative breeding schemes (i.e., $1-\frac{\frac{1}{4}\sum r_{dev}^2}{\Delta F}$, only calculated by empirical means in Chapter 2).

6.4.3. *Extensions to the stochastic simulation model for evaluating benefits from quadratic optimisation in multiple-trait breeding goals*

It is envisaged that the empirical framework developed in this thesis has good scope for being utilised in practical situations for the evaluation of the benefits from quadratic optimisation in the context of multiple-trait breeding goals when identified QTL are segregating.

From a modelling standpoint, natural extensions of the current model to cover a broader variety of scenarios would be the introduction of: i) interactions between alleles (recessive, dominant or overdominant alleles), ii) accounting for pleiotropic QTL with effects on more than one trait in the breeding goal, and iii) epistatic interactions between more than one QTL segregating in the breeding goal. Also, the model can be extended to allow for overlapping generations using available dynamic selection algorithms (Sonesson and Meuwissen, 1998 and Grundy *et al.* 2000).

A desired gains approach (Brascamp, 1984) would be a straightforward way for dealing with polygenic responses in the undesired direction due to unfavourable genetic correlations (i.e., negative in the scenario analysed in Chapter 5) between traits in the breeding goal. Including a pre-determined response value would help to prevent exacerbated extra gains in polygenic response in the undesired direction such those observed from quadratic optimisation in trait 2.

An interesting extension of the current multi-trait breeding model would be to use optimal QTL weights rather than fixed and equal weights for both the polygenic and QTL components of the total breeding value. Under single trait scenarios, optimal QTL weights in the breeding goal have been shown to be effective in maximising cumulated genetic gain over a planning horizon and to avoid early losses in polygenic gain during allele segregation (Dekkers and van Arendonk, 1998). The immediate benefit for the two-trait breeding goal analysed here would be on reducing the observed polygenic loss in the trait with greater weight in the breeding goal. A deterministic framework is available for calculating optimal QTL weights for additive and non-additive QTL and for multiple QTL (Dekkers and van Arendonk, 1998; Manfredi *et al.*, 1998; Dekkers and Chakraborty, 2001; Chackraborty *et al.*, 2002). The approach of Villanueva *et al.* (2002a) for combining both quadratic

optimisation and optimal weights for the QTL can be readily extended to when an identified QTL segregates in a multiple breeding goal.

Optimal QTL weights can also be found by stochastic means through evolutionary computation (Li *et al.*, 2001). This approach could be particularly interesting when the optimisation problem becomes too complex to be solved by deterministic means, for instance in the case of multiple traits, multiple QTL, epistatic interactions between alleles of different QTL and pleiotropic QTL.

The objective of a combined optimisation of contributions and QTL weights is to provide a general framework for maximising the gain in the overall breeding goal while using all the available information for the polygenic and QTL components of the total breeding value of all traits in the breeding goal.

6.5. General conclusion

Practical tools for the design and operation of livestock breeding schemes for maximising the rate of genetic gain while constraining inbreeding have been successfully developed and applied in this thesis. Breeders have now the possibility of designing and operating livestock breeding schemes for maximising gain while explicitly managing the risk associated with inbreeding at the same time. It has been shown that at the same rate of inbreeding, commercial populations can only benefit from applying quadratic indices. For the first time, a self-contained framework for predicting the rate of gain under quadratic optimisation is now available. Potential benefits over selection based on linear indices (i.e., truncation selection) can now be quantified *a priori*. Benefits have been predicted for a broad range of parameters and relevantly, were maintained irrespective of the size of the scheme. The key finding that in breeding schemes using quadratic indices, the Mendelian sampling term is the selective advantage not only allowed a deeper understanding of the mechanics of quadratic optimisation but also led to the relevant development of framework for predictions of the accuracy of the Mendelian sampling term. It has been also shown that quadratic optimisation has also potential for providing extra gains when

selection is on multiple-trait breeding goals in which an identified QTL is segregating in the population. This provides a practical way of assessing alternatives for including identified QTL affecting the overall breeding goal in selection decisions in breeding schemes operated by quadratic indices.

BIBLIOGRAPHY

Agerholm, J.S., Bendixen, C., Andersen, O., Arnbjerg J., 2001. Complex vertebral malformation in Holstein calves. *Journal of Veterinary Diagnosis Investigation*, **13**: 283-289.

Amer, P., Crump, R., and Simm, G., 1998. A terminal sire selection index for UK beef cattle. *Animal Science*, **67**: 445-454.

Abdel-Azim, G., and Freeman A.E., 2002. Superiority of QTL-assisted selection in dairy cattle breeding schemes. *Journal of Dairy Science*, **85**: 1869-1880.

Bijma, P., 2000. *Long-term genetic contributions - Prediction of rates of inbreeding and genetic gain in selected populations*. Dissertation, Wageningen University, Universal Press, Veenendaal, The Netherlands.

Bijma, P., and Woolliams, J.A., 1999. Prediction of genetic contributions and generation intervals in populations with overlapping generations under selection. *Genetics*, **151**: 1197-1210.

Bijma, P., and Woolliams, J.A., 2000. Prediction of rates of inbreeding in populations selected on Best Linear Unbiased Prediction of breeding value. *Genetics*, **156**: 361-373.

Bijma, P., and Rutten, M.J.M., 2002. Sel*A*ction: Software for optimisation of breeding programs. In: *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*, Montpellier, France. CD-ROM communication no.28-15.

Bijma, P., van Arendonk, J.A.M., and Woolliams, J.A.W., 2001. Predicting rates of inbreeding for livestock improvement schemes. *Journal of Animal Science*, **79**: 840-853.

Bondesson, L., 1989. Mathematical principles of optimal utilisation of clones under diversity constraints. *Swedish University of Agricultural Sciences, Section of Forest Biometry, Report 27.*

Brascamp, E.W., 1984. Selection indices with constraints. *Animal Breeding Abstracts*, **52**: 645-654.

Bulmer, M.G., 1971. The effect of selection on genetic variability. *American Naturalist*, **105**: 201-211.

Brisbane, J.R., Gibson, J.P., 1995. Balancing selection response and rate of inbreeding by including genetic relationships in selection decisions. *Theoretical Applied Genetics*. **91**: 421-431.

Burrow, H.M., 1993. The effects of inbreeding in beef cattle. *Animal Breeding Abstracts*, **61**: 737-751.

Caballero, A., Santiago, E., and Toro, M.A., 1996. Systems of mating to reduce inbreeding in selected populations. *Animal Science*, **62**: 431-442.

Caballero, A., and Santiago, E., 1998. Survival rates of major genes in selection programmes. In: *Proceedings of the 6th World Congress on Genetics Applied to Livestock Production*, Armidale, Australia, **26**: 6-12.

Caballero, A., and Toro, M.A., 2000. Interrelations between effective population size and other pedigree tools for the management of conserved populations. *Genetical Research, Cambridge*, **75**: 331-343.

Chakraborty, R., Moreau, L., and Dekkers, J.C.M., 2002. A general method to optimize selection on multiple identified quantitative trait loci. *Genetics, Selection, Evolution*, **34**: 145-170.

Correnti, G., 2002. *Optimisation with constraints in animal breeding*. M.Sc. thesis, School of Computing, Napier University, Edinburgh, UK.

de Koning, G.J., Weller, J.I., 1994. Efficiency of direct selection on quantitative trait loci for a two-trait breeding objective. *Theoretical Applied Genetics*, **88**: 669-677.

Dekkers, J.C.M., and van Arendonk, J.A.M., 1998. Optimizing selection for quantitative traits with information on an identified locus in outbred populations. *Genetical, Research, Cambridge*, **71**: 257-275.

Dekkers, J.C.M., and Chakraborty, R., 2001. Potential gain from optimizing multigenerational selection on an identified quantitative trait locus. *Journal of Animal Science*, **79**: 2975-2990.

Dempfle, L., 1975. A note on increasing the limit of selection through selection within families. *Genetical Research, Cambridge*. **24**: 127-135.

Eding, H., Crooijmans R.P.M.A., Groenen, M.A.M., Meuwissen, T.H.E., 2002. Assessing the contribution of breeds to genetic diversity in conservation schemes. *Genetics, Selection, Evolution*, 34: 613-633.

Falconer, D.S., and Mackay, T.F.C., 1996. *Introduction to Quantitative Genetics*, Fourth Edition. Longman, Harlow, UK.

F.A.O., 1998. *Management of small populations at risk. Secondary guidelines for the development of national farm genetic resources management plans*. Food and Agriculture Organization of the United Nations, Rome, Italy.

Fernández, A., Rodrigañez J., Toro, M.A., Rodriguez, M.C., Silió, L., 2002. Inbreeding effects on the parameters of the growth function in three strains of Iberian pigs. *Journal of Animal Science*, **80**: 2267-2275.

Fernando R.L., and Grossman M., 1989. Marker assisted selection using best linear unbiased prediction, Genetics, Selection, Evolution, **21**: 467-477.

Frankham, R., Ballou, J.D., and Briscoe, D.A., 2002. *Introduction to conservation genetics*. Cambridge University Press, Cambridge, UK.

Freyer, G., Sørensen, P., Kühn, C., Weikard, R., and Hoeschele I., 2003. Search for pleiotropic QTL on chromosome BTA6 affecting yield traits of milk production. *Journal of Dairy Science*, **86**: 999-1008.

Gibson, J.P., 1994. Short-term gain at the expense of long-term response with selection on identified loci. In: *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production*, Guelph, Canada. **21**: 201-204.

Goddard, M.E., 1992. Optimal effective population size for the global population of black and white dairy cattle. *Journal of Dairy Science*, **75**: 2902-2911.

Goddard, M.E., 2003. Animal breeding in the (post-) genomic era. *Animal Science*, **76**: 353-365.

Goddard, M.G., and Smith, C., 1990. Optimum number of bull sires in dairy cattle breeding. *Journal of Dairy Science*, **73**: 1113-1122.

Goddard, M.E., and Hayes, B.J., 2002. Optimisation of response using molecular data. In: *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*, Montpellier, France. CD-ROM communication no.22-01.

Grisart, B., Coppieters, W., Farnir, F., Karim, L., Ford, C., Berzi, P., Cambisano, N., Mni, M., Reid, S., Simon, P., Spelman, R., Georges, M., and Snell, R., 2002. Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. *Genome Research*, **12**: 222-231.

Groeneveld, E., Kovac, M., and Wang, T., 1990. PEST, a general purpose BLUP package for multivariate prediction and estimation. In: *Proceedings of the 4th World Congress on Genetics Applied to Livestock Production*, Edinburgh, Scotland, **8**: 488-491.

Grundy, B., Caballero, A., Santiago, E., and Hill, W.G., 1994. A note on using biased parameter values and non-random mating to reduce rates of inbreeding in selection programmes. *Animal Production*, **59**: 465-468.

Grundy, B., Villanueva, B., and Woolliams, J.A., 1998a. Dynamic selection procedures for constrained inbreeding and their consequences for pedigree development. *Genetical Research, Cambridge*, **72**: 159-168.

Grundy, B., Luo, Z. W., Villanueva, B., and Woolliams, J.A., 1998b. The use of Mendelian indices to reduce the rate of inbreeding in selection programmes. *Journal of Animal Breeding and Genetics*, **115**: 39-51.

Grundy, B., Villanueva, B., and Woolliams, J.A., 2000. Dynamic selection for maximising response with constrained inbreeding in schemes with overlapping generations. *Animal Science*, **70**: 373-382.

Gutiérrez, J.P., Altarriba, J., Díaz C., Quintanilla R., Cañon, J., and Piedrafita, J., 2003. Pedigree analysis of eight Spanish beef cattle breeds. *Genetics, Selection, Evolution*, 35: 1-21.

Hall, S.J.G., and Hall, J.G., 1988. Inbreeding and population dynamics of the Chillingham cattle (*Bos taurus*). *Journal of Zoology, London*, **216**: 479-493.

Hill, W.G., 1974. Prediction and evaluation of response to selection with overlapping generations. *Animal Production*, **18**: 117-139.

Hill, W.G., 2000. Maintainance of quantitative genetic variation in animal breeding programmes. *Livestock Production Science*, **63**: 99-109.

Hill, W.G., Caballero, A., Dempfle, L., 1996. Prediction of response to selection within families. *Genetics, Selection, Evolution*, **28**: 379-383.

Hill, W.G., and Zhang X.-S., 2002. Genetic variation within and among animal populations. In: *Simm, G., B. Villanueva, K.D. Sinclair, and S. Townsend (ed.) Farm Animal Genetic Resources. BSAS Occasional Publication No. 30.* Nottingham University Press, Nottingham, UK (*In press*).

Keller, D.S., Gearheart, W.W., Smith, C., 1990. A comparison of factors reducing selection response in closed nucleus breeding schemes. *Journal of Animal Science*, **68**: 1553-1561.

Kennedy, B.W., Quinton, M., van Arendonk, J.A.M., 1992. Estimation of effects of single genes on quantitative traits. *Journal of Animal Science*, **70**: 2000-2012.

Kinghorn, B.P., 1998. Application of reproductive technologies. In: *Proceedings of the 6th World Congress on Genetics Applied to Livestock Production*, Armidale, Australia, Production genetic field day: 36-39.

Kinghorn, B.P., Meszaros, S.A., and Vagg, R.D., 2002. Dynamical tactical decision systems for animal breeding. In: *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*, Montpellier, France. CD-ROM communication no.23-07.

Lacy, R.C., and Ballou, J.D., 1998. Effectiveness of selection in reducing the genetic load in populations of *Peromyscus polionotus* during generations of inbreeding. *Evolution*, **52**: 900-909.

Lande R., and Thompson, R., 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics*, **124**: 743-756.

Lamberson, W.R., and Thomas, D.L., 1984. Effects of inbreeding in sheep: a review. *Animal Breeding Abstracts*, **52**: 287-297.

Lewis, R.M., and Simm, G., 2000. Selection strategies in sire referencing schemes in sheep. *Livestock Production Science*, **67**: 129-141.

Li, Y., van der Werf, J.H.J., and Kinghorn, B.P., 2001. Optimization of selection for two identified quantitative trait loci. In: *Proceedings of the Association of Advancement of Animal Breeding and Genetics*, **14**: 265-268.

Lynch, M. and Walsh, B., 1998. *Genetic analysis of quantitative traits*. First Edition, Sinauer Associates, Sunderland.

Maroo, S., 1999. *A genetic analysis of the Exmoor pony population*. M.Sc. Thesis. Edinburgh University, UK

Manfredi E., Barbieri, M., Fournet, F., Eisen, J.M., 1998. A dynamic deterministic model to evaluate breeding strategies under mixed inheritance. *Genetics, Selection, Evolution*, **30**: 127-148.

Meuwissen, T.H.E., 1997. Maximising the response of selection with predefined rate of inbreeding. *Journal Animal Science*, **75**: 934-940.

Meuwissen, T.H.E., and Luo, Z., 1992. Computing inbreeding coefficient in large populations. *Genetics, Selection, Evolution*, **24**: 305-313.

Meuwissen, T.H.E., and Woolliams, J.A., 1994a. Effective sizes of populations to prevent a decline in fitness. *Theoretical Applied Genetics*, **89**: 1019-1026.

Meuwissen, T.H.E., and Woolliams, J.A., 1994b. Response versus risk in breeding schemes. In: *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production*, Guelph, Canada. **18**: 236-243.

Meuwissen, T.H.E., and Woolliams, J.A., 1994c. Maximizing genetic response in breeding schemes of dairy cattle with constraint on variance of response. *Journal of Dairy Science*, **77**: 1905-1916.

Meuwissen, T.H.E., and Sonesson, A. K., 1998. Maximizing the response of selection with predefined rate of inbreeding: Overlapping generations. *Journal of Animal Science*, **76**: 2575-2583.

Meszaros, S.A., Banks, R.G., van der Werf J.H.J., and M.E. Goddard., 1999. The use of genetic algorithms for optimising age structure in breeding populations when inbreeding depresses genetic gain through effects on reproduction. *Animal Science*, **68**: 449-457.

M.L.C., 1999. *Sheep yearbook*. Meat and Livestock Commission, Milton Keynes, UK.

Navarro, P., Koerhuis, A.N.M., Chatziplis, D., Visscher, P.M., and Haley, C.S., 2002. Genetics studies of ascites in a broiler population. In: *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*, Montpellier, France. CD-ROM communication no.04-14.

Nicholas, F.W., 1987. *Veterinary Genetics*. Oxford University Press, New York, US.

Nomura, T., Honda, T., and Mukai, F., 2001. Inbreeding and effective population size of Japanese Black Cattle. *Journal of Animal Science*, **79**: 366-370.

Notter, D.R., 1999. The importance of genetic populations diversity in livestock populations of the future. *Journal of Animal Science*, **77**: 61-69.

Pong-Wong, R., and Woolliams, J.A., 1998. Response to mass selection when an identified major gene is segregating. *Genetics, Selection, Evolution*, **30**: 313-337.

Press, W., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P., 1992. *Numerical recipes in FORTRAN: the art of scientific computing*, Second edition. Cambridge University Press.

Pryce, J.E., Coffey, M.P., Brotherstone, S.H., and Woolliams, J.A., 2002. Genetic relationships between calving interval and body condition score conditional on milk yield. Journal of Dairy Science, **85**: 1590-1595.

Quinton, M., Smith, C., and Goddard, M.E., 1992. Comparison of selection methods at the same level of inbreeding. *Journal Animal Science*, **70**: 1060-1067.

Rodrigañez, J., Toro, M.A., Rodriguez, M.C., Silió, L., 1998. Effect of founder allele survival and inbreeding depression on litter size in a close line of Large White pigs. *Animal Science*, **67**: 573-582.

Robertson, A., 1960. A theory of limits in artificial selection. *Proceedings of the Real Society of London, Series B* **153**: 234-249.

Roberston, A., 1965. The interpretation of genotypic ratios in domestic animal populations. Animal Production, **7**: 319-324.

Ruane J., and Colleau J.J., 1995. Marker assisted selection for genetic improvement of animal populations when a single QTL is marked. *Genetical Research, Cambridge*, **66**: 71-83.

Ruane J., and Colleau J.J., 1996. Marker assisted selection for a sex-limited character in a nucleus breeding population, *Journal of Dairy Science*, **79**: 1666-1678.

Rutten, M.J.M, Bijma, P., Woolliams, J.A., and van Arendonk, J.A.M., 2002. SelAction: Software to predict selection response and rate of inbreeding in livestock breeding programs. *The Journal of Heredity*, 93: 456-458.

Sánchez, L., Bijma, P., and Woolliams, J.A., 2003. Minimizing inbreeding by managing genetic contributions across generations. *Genetics*, **164**: 1589-1595.

Santiago, E., and Caballero, E., 1995. Effective size of populations under selection. *Genetics*, **139**: 1013-1030.

Schoenauer, M., and Michalewicz, Z., 1997. Evolutionary computation. *Control and Cybernetics*, 26: 307-338.

Schulman, N.F., De Vries, M.J., and Dentine, M.R., 1999. Linkage disequilibrium in two-stage marker-assisted selection. *Journal of Animal Breeding and Genetics*, **116**: 99-110.

Shepherd, R.K., and Kinghorn, B.P., 1998. A tactical approach to the design of crossbreeding programs. In: *Proceedings of the 6th World Congress on Genetics Applied to Livestock Production*, Armidale, Australia, **25**: 431-438.

Short, R.E., MacNeil, M.D., Grosz, M.D., Gerrard, D.E., and Grings, E.E., 2002, Pleiotropic effects in Hereford, Limousin, and Piedmontese F_2 crossbred calves of genes controlling muscularity including the Piedmontese myostatin allele. *Journal of Animal Science*, **80**: 1-11.

Simm, G., 1998. *Genetic improvement of cattle and sheep*. Farming Press, Ipswich, UK.

Simm, G., and Dingwall, W.S., 1989. Selection indices for lean meat production in sheep. *Livestock Production Science*, **21**: 223-233.

Smith, L.A., Casell, B.G., and Pearson, R.E., 1998. The effects of inbreeding on the lifetime performance of dairy cattle. *Journal of Dairy Science*, **81**: 2729-2737.

Sonesson, A.K., and Meuwissen, T.H.E., 2000. Mating schemes for optimum contribution selection with constrained rates of inbreeding. *Genetics, Selection, Evolution*, **32**: 231-248.

Sonesson, A.K., and Meuwissen, T.H.E., 2002. Non-random mating for selection with restricted rates of inbreeding and overlapping generations. *Genetics, Selection, Evolution*, **34**: 23-39.

Sonesson, A.K., Grundy, B., J.A., Woolliams, J.A.W., and Meuwissen, T.H.E., 2000. Selection with control of inbreeding in populations with overlapping generations: a comparison of methods. *Animal Science*, **70**: 1-8.

Sonesson, A.K., Janss, L.G., Meuwissen, T.H.E., 2003. Selection against genetic defects in conservation schemes while controlling inbreeding. *Genetics, Selection, Evolution*, **35**: 353-368.

Thibier, M., and H.-G., Wagner., 2002. World statistics for artificial insemination in cattle. *Livestock Production Science*, **74**: 203-212.

Thompson, J.R., Everett, R.W., and Hammerschmidt N.L., 2000. Effects of inbreeding on production and survival in holsteins. *Journal of Dairy Science*, **83**: 1856-1864.

Toro, M.A., and Nieto, B., 1984. A simple method for increasing the response to artificial selection. *Genetical Research, Cambridge*, **69**: 145-158.

Toro, M.A., Nieto, B., and Salgado, C., 1988. A note on minimization of inbreeding in small-scale selection programmes. *Livestock Production Science*, **20**: 317-323.

Toro, M.A., and Pérez-Enciso, M., 1990. Optimization of selection response under restricted inbreeding. *Genetics, Selection, Evolution,* **22**: 93-107.

Verrier, E., Colleau, J.J., Foulley, J.L., 1993. Long-term effects of selection based on the animal model BLUP in a finite population. *Theoretical Applied Genetics*, **87**: 446-454.

van Arendonk, J.A.M, and Bijma, P., 2002. Utilizing reproduction techniques in animal breeding programmes. In: *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*, Montpellier, France. CD-ROM communication no.27-02.

Verrier, E., 2001. Marker assisted selection for the improvement of two antagonistic traits under mixed inheritance. Genetics, Selection, Evolution, **33**: 17-38.

Villanueva, B., Woolliams, J.A., Simm, G., 1994. Strategies for controlling rates of inbreeding in MOET nucleus schemes for beef cattle. *Genetics, Selection, Evolution*, **26**: 517-535.

Villanueva B., G. Simm, and, Woolliams, J.A., 1995. Genetic progress and inbreeding for alternative nucleus breeding schemes for beef cattle. *Animal Science*, **61**: 231-239.

Villanueva, B., Woolliams, J.A., and Gjerde, B., 1996. Optimum designs for breeding programs under mass selection with an application in fish breeding. *Animal Science*, **63**: 563-576.

Villanueva, B. and Woolliams, J.A., 1997. Optimization of breeding programs under index selection and constrained inbreeding. *Genetical Research, Cambridge*, **69**: 145-158.

Villanueva, B., Pong-Wong, R., and Woolliams, J.A., 1999. Potential benefit from using and identified major gene and BLUP evaluation with truncation and optimal selection. *Genetics, Selection, Evolution*, **34**: 679-703.

Villanueva, B., Bijma, P., Woolliams, J.A., 2000. Optimal mass selection policies for schemes with overlapping generations and restricted inbreeding. *Genetics, Selection, Evolution*, **32**: 339-355.

Villanueva, B., Dekkers, J.C.M., Woolliams, J.A., Settar P., 2002a. Maximising genetic gain with QTL information and control of inbreeding. In: *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*, Montpellier, France. CD-ROM communication no.22-18.

Villanueva, B., Pong-Wong, R., and Woolliams, J.A., 2002b. Potential benefit from using an identified major gene and BLUP evaluation with truncation and optimal selection. *Genetics, Selection, Evolution*, **31**:115-133.

Villanueva, B., Pong-Wong, R., Woolliams, J.A.W., and Avendaño, S. 2003. Managing genetic resources in selected and conserved populations. In: *Simm, G., B. Villanueva, K.D. Sinclair, and S. Townsend (ed.) Farm Animal Genetic Resources. BSAS Occasional Publication No. 30.* Nottingham University Press, Nottingham, UK (*In press*).

Visscher, P.M., Smith, D., Hall, S.J.G., Williams, J.A., 2001. A viable herd of genetically uniform cattle. *Nature*, **409**: 303.

Weigel, K.A., 2001. Controlling inbreeding in modern breeding programs. *Journal of Dairy Science*, **84**(E.Suppl.): E177-E184.

Weigel, K.A., and Lin, S.W., 2002. Controlling inbreeding by constraining the average relationship between parents of young bulls entering AI progeny test programs. *Journal Dairy Science*, **85**: 2376-2383.

Wiener, G., Lee, G.J., Woolliams, J.A., 1994. Consequences of inbreeding for financial returns from sheep. *Animal Production*, **59**: 245-249.

Wiener, P., Smith, A.L., Lewis, A.M., Woolliams, J.A., Williams, J.L., 2002. Muscle-related traits in cattle: The role of the myostatin gene in the South Devon breed. *Genetics, Selection, Evolution*, **34**: 221-232.

Woolliams, J.A., 1989. Modifications to MOET nucleus breeding schemes to improve rates of genetic gain and decrease rates of inbreeding in dairy cattle. *Animal Production*, **49**:1-14.

Woolliams, J.A., and Meuwissen, T.H.E., 1993. Decision rules and variance of response in breeding schemes. *Animal Production*, **56**: 179-186.

Woolliams, J.A. and Thompson, R., 1994. A theory of genetic contributions. In: *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production*, Guelph, Canada. **19**: 127-134.

Woolliams, J.A., and Bijma, P., 2000. Predicting rates of inbreeding in populations undergoing selection. *Genetics*, **154**: 1851-1864.

Woolliams, J.A., and Mäntysaari, E.A., 1995. Genetic contributions of Finnish Ayrshire bulls over four generations. *Animal Science*, **61**: 177-187.

Woolliams, J.A., Bijma, P., and Villanueva, B., 1999. Expected genetic contributions and their impact on gene flow and genetic gain. *Genetics*, **153**: 1009-1020.

Woolliams, J.A., Pong-Wong, R., and Villanueva, B. 2002. Strategic optimisation of short- and long-term gain and inbreeding in MAS and non-MAS schemes. In: *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*, Montpellier, France. CD-ROM communication no.32-02.

Wray, N.R., Hill, W.G., 1989. Asymptotic rates of response from index selection. *Animal Production*, **49**: 217-227.

Wray, N.R., and Thompson, R., 1990. Prediction of rates of inbreeding in selected populations. *Genetical Research, Cambridge*. **72**: 41-54.

Wray, N.R., Woolliams, J.A., and Thompson, R., 1990. Methods for predicting rates of inbreeding in selected populations. *Theoretical Applied Genetics*, **80**: 503-512.

Wray, N.R., and Simm, G., 1991. *BLUP for pedigree beef and sheep breeders*. SAC Technical Note T265. The Scottish Agricultural College, West Mains Road, Edinburgh, UK.

Wray, N.R., and Goddard, M.E., 1994. Increasing long-term response to selection. *Genetics, Selection, Evolution*, **26**: 431-451.

Xie, C., and Xu, S., 1998. Efficiency of multistage marker assisted selection in the improvement of multiple traits. *Heredity*, **80**: 489-498.